

Theoretical Design and Performance Assessment of a BSO-Based Optical Correlator

Timothy George Slack

**A thesis submitted for the degree of:
DOCTOR OF PHILOSOPHY**

Submitted to:

**THE LASER & OPTICAL SYSTEMS ENGINEERING GROUP;
THE DEPARTMENT OF MECHANICAL ENGINEERING;**

at

THE UNIVERSITY OF GLASGOW;

in

September 1993.

©T.G.Slack, 1993

**British Aerospace,
Sowerby Research Centre,
Bristol.**

ProQuest Number: 13818521

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13818521

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Thesis
9491
copy 1



Synopsis

The main objective of the research, presented in this thesis, was the design of an updatable optical correlator. This was based on the use of a crystal of Bismuth Silicon Oxide (BSO) placed in the Fourier plane. The system performance was prescribed to meet specific targets: the correlation of images comprising 512×512 pixels; at 40 milliseconds per image. To undertake the design work, the predictions of the Band Transport Model were used to describe the performance of the BSO crystal. In addition to this, coupled wave theory was used to model the effects of a 'thick' hologram in the correlator.

Several, previously unreported, aspects of the system were considered in the design. These included the effects of the positions of the input images, with respect to the transforming lenses. It was shown theoretically, that spatial filtering effects occur if the images are not in the front focal planes of the lenses. Specification of the lens combinations to be used in the correlator was carried out using standard ray-tracing software with manufacturer's lens data. Digital simulation was also used to predict intensities in both the Fourier and the correlation plane of the optical system. An important effect, not hitherto reported, was that of the hologram's grating vector becoming tilted with respect to the electric field which is applied to the BSO. The drop in diffraction efficiency thus caused can be comparable with that due to the Bragg de-tuning of the beam reading out the hologram.

It was shown that the correlator could be designed to meet the imposed specifications. The main problems encountered were the angular selectivity of the BSO hologram, and low writing-beam intensities in the Fourier plane. In order to predict the maximum sizes of the images that could be processed, the equations governing system performance were expressed in terms of a reduced parameter set. Limits on the number of pixels, in the images on the read and write sides of the correlator, were then determined.

It was predicted that it would be difficult to correlate two images of more than about 2000 pixels across. This limit was imposed by the angular selectivity effects. It

was also predicted that a system would be feasible, which could correlate an image of 6000×6000 pixels with a small reference template. In both cases, optical path lengths of several meters would be needed; as well as large-aperture, thin BSO crystals. From the results it was concluded that BSO was not going to compete with other, recently developed, materials and devices: such as switchable photo-chromic materials and phase modulating spatial light modulators (SLMs).

Severe problems with the acquisition of suitable SLMs, meant that the designed correlator could not be constructed. Effort was therefore diverted to extend the digital simulation work to cover these devices. Models were developed which represented the distortions typically found in optically addressed SLMs. The results produced highlighted some shortcomings of published work in this field. Finally, an experimental method was chosen, which allowed combinations of different distortions to be applied to pairs of images. This method, and the computer models developed, can be used to assess fully the effects of SLM distortions on any optical processing task. An example is presented in which correlations were performed in the presence of a grey-level non-uniformity and two types of phase distortion. The effects of the distortions on peak-height, signal-to-noise ratio, shift-invariance and the accuracy of peak location were computed.

Acknowledgements

I would like to thank both my external supervisor, Dr Chris Chatwin, and my internal supervisor, Nigel Aldridge, for their help and support during the undertaking of this research. I would also like to thank British Aerospace plc. for their financial support.

Contents

1	Introduction	15
1.1	Optical Correlators	17
1.1.1	Optical vs Digital Correlation	17
1.1.2	Optical Correlation: Methods and Uses	19
1.1.3	A History of BSO-Based Correlators	23
1.1.4	Spatial Light Modulators and Input Images	30
1.2	The Photorefractive Effect and the Band Transport Model	33
1.3	The Objectives and Scope of this Research	37
2	Thick Holograms in an Optical Correlator	47
2.1	Maximum Diffraction Efficiency	50
2.1.1	Refractive Index Modulation of BSO	52
2.2	Variations in Diffraction Efficiency	56
2.2.1	Changes in Read-Beam Angle	57
2.2.2	Changes in Write-Beam Angle	60
2.2.3	Beam Angles Produced in the Correlator	66
2.3	Validity of the Use of Coupled Wave Theory	68
2.4	Depth of Field in the Fourier Plane	75
2.5	Conclusions	76
3	Other Factors Pertinent to the Correlator	80
3.1	Effect of Input Plane Positions	84
3.1.1	The Location of the Correlation Plane	84
3.1.2	Unwanted Spatial Filtering Effects	90

3.2	Lens and Wavelength Requirements	96
3.2.1	Two-Element Compound Lenses	97
3.3	Powers Required in the Read and Write-Beams	108
3.3.1	Correlation Plane Intensities	109
3.3.2	Fourier Plane Intensities	113
3.3.3	Response Time of the BSO	115
3.4	Tilted Fringes in an Applied Field	119
3.5	Limits on Fringe Spacing	122
3.6	Comparison of BSO and Holographic Film	124
3.7	Beam Polarisation	135
3.8	Conclusions	136
4	Digital Simulation of SLM Distortions	142
4.1	Image Processing Methods	144
4.1.1	Non-Linear Grey-Level Transfer	144
4.1.2	Random Grey-Level Noise	149
4.1.3	Structured Grey-Level Distortions	152
4.1.4	Large Scale Phase Distortions	152
4.1.5	Small Scale Phase Noise	161
4.1.6	Signal Dependent Phase Distortion	164
4.1.7	Device Resolution	167
4.2	A 2^N Factorial Experiment	168
4.2.1	The Distortions Used in the Experiment	170
4.2.2	Experimental Results	172
4.2.3	Discussion of the Results	176
4.3	Conclusions	179
5	The Design of the Optical Correlator	183
5.1	Calculations for the Specified Optical System	186
5.1.1	Steps used in the Definition	188
5.1.2	Calculated values for the System	194

5.2	Reducing the Set of Variable Parameters	198
5.3	Extending the Correlator to Larger Images	204
5.3.1	Maximum Image Size on the Read-Side	205
5.3.2	Large Image on the Write-Side	207
5.4	Conclusions	211
5.4.1	Discussion of the Optical System Design	212
5.4.2	Discussion of the Processing of Larger Images	214
6	Conclusions and Recommendations	216
6.1	Design Meeting Required Specifications	217
6.2	Extending the Design to Larger Images	218
6.3	Digital Simulation of SLM Performance	220
6.4	Recommendations for Future Work	221
A	Operations Required to Perform Correlation	224

List of Figures

1.1	Reference image (left), input image and correlation planes without edge enhancement (producing a failed result) and with a simple edge enhancing filter giving a good recognition peak.	18
1.2	A comparison between the Vander Lugt Correlator (top diagram) and the Joint Transform Correlator (bottom diagram).	23
1.3	Images of machined component (left) with simple computer generated hole template inset and their computed cross-correlation (right). . . .	31
1.4	Grating formed by charge separation caused by cosinusoidal illumination.	36
2.1	Geometry used to describe the diffraction from the volume hologram.	51
2.2	$ E_{SC} $ varying with fringe spacing for various applied fields.	54
2.3	Coupling in two wave mixing as a function of fringe spacing. Circles show experimental data while the solid line shows a normalised theoretical curve for $\Lambda_C=0.55\mu\text{m}$	56
2.4	Diffraction efficiency vs read-beam angle at a constant fringe spacing but for different thicknesses of grating.	58
2.5	Diffraction efficiency vs read-beam angle for different fringe spacings but with a constant thickness of $d = 1\text{mm}$	58
2.6	Grating formed by two interfering beams.	60
2.7	Varying the writing object beam angle with fixed reference and read-beams.	61
2.8	Diffraction efficiency varying with object beam angle for a system initially set up with $\Lambda = 7\mu\text{m}$	63

2.9	Diffraction vs object beam angle for a 1mm thick hologram with an initial fringe spacing of $2\mu\text{m}$	65
2.10	Fourier transform of a grating image.	66
2.11	Fringe profile for different grating intensity modulations.	71
2.12	Diffraction efficiency varying with fringe spacing for a 1mm thick BSO hologram with an applied field of 5kVcm^{-1} . Diffraction efficiency should be constant (solid line).	72
2.13	Experimental data (circles) and theoretical prediction (curve) for variation of diffraction efficiency with read-beam angle.	75
3.1	Schematic diagram of reading and writing arrangements and the correlation plane.	85
3.2	Fourier transforming a Dirac delta function	86
3.3	Simple holographic lens replayed by plane wave	87
3.4	System formed of lens f_r and holographic lens	88
3.5	Variation of diffraction efficiency with read-image spatial frequency for points at 0, 0.01 and $0.02f$ from the optic axis. The input plane is $0.75f$ from the Fourier transforming lens.	92
3.6	Variation of diffraction efficiency with read-image spatial frequency for points at 0, 0.01 and $0.02f$ from the optic axis. The input plane is $0.25f$ from the Fourier transforming lens.	93
3.7	Variation of diffraction efficiency with write-image spatial frequency for points at 0, 0.01 and $0.02f$ from the optic axis. The input plane is $0.25f$ from the Fourier transforming lens.	93
3.8	Slices through three computer generated correlation planes with a 'perfect' result on the left and progressive degrees of defocussing in the two other peaks. The BSO hologram was simulated with a reference/d.c. beam ratio of 1:4.	95
3.9	Repeat of the above but with a reference/d.c. beam ratio of 1:32. . .	95
3.10	A two element compound lens.	97
3.11	Thick lens version of the two-element compound lens.	99

3.12	Wavefront aberrations produced by a lens.	102
3.13	Aberrated spot diagrams with schematic representation of the diffraction limited spot size. Top is a 452mm focal length lens formed from a 200mm doublet followed by a -300mm biconcave lens. Bottom is a single 400mm doublet for comparison.	106
3.14	Spot diagrams for a 452mm focal length lens in which a -300mm biconcave lens is preceded by a 200mm doublet.	107
3.15	The variation of the fundamental component of space charge field with the normalised variable t/τ_g	116
3.16	Decay time versus erase beam intensity for grating spacings of 10 microns (circles) and 22 microns (triangles).	118
3.17	Decay time versus, E_A/E_C with grating spacings of 10 microns (triangles) and 22 microns (circles): wash-out beam intensity of 3mWcm^{-2} . Theoretical curves generated with $\Theta^2 = 160$ and $\tau_c = 5.2$	118
3.18	Decay time versus grating spacing (circles) for $E_A=7\text{kVcm}^{-1}$: wash-out beam intensity of 3mWcm^{-2} . Theoretical curve generated with $\Theta^2 = 120$ and $\tau_c = 5.0$	119
3.19	Fringes tilted with respect to the applied field.	119
3.20	Diffraction efficiency versus fringe tilt for a 10 micron grating with an applied field of 7kV. The zero degree point corresponds to a 0 degree angle between grating vector and applied field direction.	120
3.21	Schematic diagram used to determine the fringe tilts produced at the BSO hologram by rays from different parts of the write-image.	121
3.22	Hypothetical film response curve.	125
3.23	Sinusoidal fringes and the result of their non-linear recording.	127
3.24	Fourier transforms of the perfect sinusoidal fringes (left) and the same fringes subject to the film transfer curve (right).	128
3.25	Fourier transform of a square pulse (left) and the modified transforms recorded with the reference beam balanced to the d.c. of the transform with the amplitude of the d.c. scaled to 16, 7 and 4.5.	130

3.26	Reconstructions of the FTs from the previous figure.	130
3.27	Modified transforms recorded with the reference beam balanced to the third lobe and with the original FT amplitude scaled by factors of 10, 2.5, 1.0 and 0.5	131
3.28	Reconstructions of the various FTs from the previous figure.	132
3.29	Modified transforms recorded with the reference beam balanced to the d.c. and the third lobe using the BSO model to compute the amplitude of the hologram.	133
3.30	Reconstructions of the various FTs from the previous figure.	133
3.31	Square pulse test shape with random noise.	134
3.32	Reconstructions of the noise corrupted square pulse after simulated FT storage on holographic film (left) and BSO (right).	135
4.1	Generic grey-level transfer curve.	145
4.2	Perfect images of component#2 and component#3 (top) and two dif- ferent grey-level distorted versions of component#2 (bottom).	146
4.3	First set of transfer curves.	147
4.4	Second set of transfer curves.	147
4.5	Correlation discrimination, between components#2 and #3, varying with the different grey-level transfer curves.	148
4.6	Gaussian populations, generated by averaging values from Fortran random number generator. The dashed lines represent perfect Gaus- sian curves.	150
4.7	Noise corrupted image of component#2 (left) and the result of simu- lating its storage as a BSO hologram (right).	151
4.8	Corresponding slices through the noise-corrupted correlation plane (left) and the perfect auto-correlation (right).	151
4.9	Computed interferograms of four phase distortions with different (x, y) variations. An intensity representation of the phase is inset in the top left corners of each of the interferograms.	154
4.10	Effect of a phase ramp on the position of the Fourier plane.	155

4.11 Auto correlation of component#2: BSO correlator with 1/16 beam amplitude ratio	155
4.12 Cross correlation of undistorted image of component#2, with second image with 4π phase ramp added.	156
4.13 Equivalent effect of spherical phase distortion at the SLM.	156
4.14 Geometry for calculating the apparent focal length from the phase change between the centre and the edge of the field.	157
4.15 Cross-correlation of undistorted image of component#2, with second image with π spherical phase front added.	158
4.16 Input images used in the position invariance experiment.	160
4.17 Correlation of single, centred hole template with an array of identical objects.	160
4.18 Correlation of hole template with array of holes, with identical, cylin- drical phase distortion added to both images.	161
4.19 Comparison of mirror and LCLV fringes	162
4.20 Simulated results of edge enhancement, by BSO holographic storage, for images with (right) and without (left) phase noise.	163
4.21 Effect of amplitude-phase cross coupling on edge enhancement.	165
4.22 Input data for discrimination experiment.	166
4.23 Discrimination experiment with no phase coupling.	166
4.24 Improved discrimination as a result of phase coupling.	166
4.25 MTF curve, and the result of applying it to the image of compo- nent#2. The original image is shown on the left, the filtered image on the right.	167
4.26 Experimental system influenced by three factors each of which can be set at two levels.	168
4.27 Computer generated phase profile including small scale phase distor- tions superimposed on an overall, large scale astigmatic distortion. . .	171
4.28 Edge enhancement of undistorted component and template, together with their correlation surface.	173

4.29	Edge enhancement and correlation results when the three distortions were present at their highest levels.	173
5.1	Steps used in defining the optical correlator parameters.	189
5.2	Comparison between the Fourier transforms of the hole template (left) and the machined component (right).	193
5.3	Scale diagram of the optical correlation system.	197
5.4	η_{min}/η_{max} plotted as a function of N/r_{BSO}	202
5.5	Schematic diagram of images and correlation plane for a small refer- ence template correlated with a large image.	210

List of Tables

- 4.1 Final results; calculated from the averaged experimental results for
the different treatment combinations. 177

- 5.1 Parameters specified for the Optical System 187
- 5.2 Parameters determining the writing intensities in the Fourier plane. . 195
- 5.3 Applied fields and associated response times. 195
- 5.4 Diffraction efficiency parameters for holograms of the two images. . . 195
- 5.5 Correlation plane intensities for the auto-correlations of the two images.196

Chapter 1

Introduction

Contents of Chapter 1

1.1	Optical Correlators	17
1.1.1	Optical vs Digital Correlation	17
1.1.2	Optical Correlation: Methods and Uses	19
1.1.3	A History of BSO-Based Correlators	23
1.1.4	Spatial Light Modulators and Input Images	30
1.2	The Photorefractive Effect and the Band Transport Model	33
1.3	The Objectives and Scope of this Research	37

Abbreviations/symbols used in Chapter 1

Ar ⁺	Argon ion (laser).
BSO	Bismuth Silicon Oxide.
BTM	Band Transport Model.
CCD	Charge Coupled Device.
CHF	Circular Harmonic Filter.
CRT	Cathode Ray Tube.
d.c.	Zero-spatial-frequency component of FT.
FT	Fourier Transform.
He-Ne	Helium-Neon (laser).
JTC	Joint Transform Correlator.
MSF	Matched Spatial Filter.
SBWP	Space Bandwidth Product.
SDF	Synthetic Discrimination Function.
SLM	Spatial Light Modulator.
TV	Television.
VLC	Vander Lugt Correlator.
$N \times N$	Number of pixels in a (square) image.

The research presented herein was carried out as part of a three year, collaborative project entitled Hybrid Industrial Inspection (HII). The project as a whole sought to construct and evaluate an industrial inspection system composed of integrated digital and coherent optical processing systems. At the centre of the optical processing system was to be an updatable optical correlator based around photorefractive Bismuth Silicon Oxide (BSO). In this introduction, the history and uses of optical correlation are reviewed in order to explain its powers and drawbacks. This review leads to the requirement for an updatable (or re-programmable) system and explains the selection of BSO as the storage medium. A literature survey is then used to sketch an outline of the history of correlation work using BSO. A generic industrial inspection task is described, which was used to specify the required correlator performance. The role of real-time spatial light modulators is also discussed. In order to provide relevant background information, the hologram formation process is briefly reviewed in terms of the band transport model. Finally, the scope, objectives and organisation of the remainder of the thesis are summarised.

1.1 Optical Correlators

Correlation, or template matching, can be used to perform two related functions. Firstly, it can be used to recognise the presence of a particular object within an image containing many different shapes. Secondly, it can be used to determine the position of a particular object within an image [1.1-1.3]. Correlation can be visualised as the process of stepping one image (the reference) over another (the object image) and, at each point, multiplying the values of their overlapping pixels. At each point, the correlation plane value is calculated by summing all these multiplied values.

1.1.1 Optical vs Digital Correlation

The step by step method of forming the correlation plane can be carried out digitally. However, for large images it is computationally more efficient to perform the correlation in the Fourier domain as shown in appendix A. In this case, the Fourier transform (FT) of each image is computed and the complex conjugate of one is mul-

multiplied by the other. The inverse transform of the result is then computed to form the correlation plane ([1.4] for example). In general, the mere performance of the mathematical operation of correlation is not, in itself, useful and it is necessary to carry out a preliminary edge enhancement of the images. A simple illustration of this is shown in figure 1.1. Here, the task is to recognise and locate a relatively low intensity square while discriminating against a brighter triangle. The figure shows correlation planes produced both with and without edge enhancement of the images. The edge enhancement in this case being provided by a simple Gaussian ‘stop’ blocking the zero-frequency components (d.c.) of the Fourier transforms.

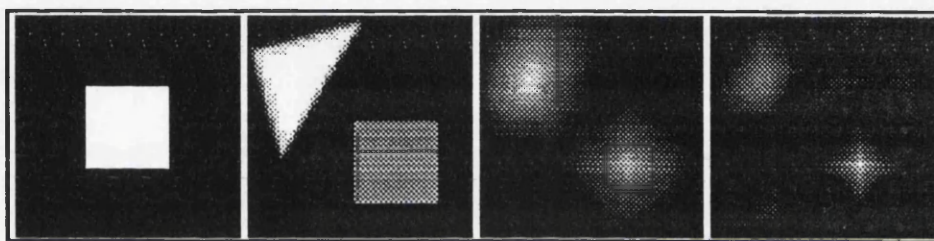


Figure 1.1: Reference image (left), input image and correlation planes without edge enhancement (producing a failed result) and with a simple edge enhancing filter giving a good recognition peak.

In order to implement this spatial filtering and correlation with two $N \times N$ pixel images, the number of essential floating point operations is shown, in appendix A, to be approximately $20 \times N^2 \log_2(N)$. For a 512 by 512 pixel image this is about 47 million operations. Since the two initial FTs can be performed in parallel, a 1.3 Gflop processor would theoretically be able to carry out this correlation at TV frame rates. However, at the time of writing, such machines cost around \$1 million and, at the same time, current CCD cameras are capable of producing images of 2048×2048 pixels. Thus, to carry out correlations, on images larger than 512×512 pixels, at TV frame rates, it is necessary to consider other available technologies in order to reduce the cost and size of the system

It was for these reasons that the Fourier transforming capability of a positive lens [1.5] generated so much interest. A further step forward was the use of holographic film to store the optically produced FTs. The stored FTs are also known as matched

spatial filters (MSFs). These MSFs could be multiplied with the transforms of other images passed through one arm of the original recording system [1.6]. Twenty eight years ago when this reference appeared the digital computing speeds available now were not even dreamed of.

1.1.2 Optical Correlation: Methods and Uses

Since the original work, many papers have been published on optical processing by researchers all around the world. In its infancy, this work made use of holographic filters stored on high resolution photographic plate. The necessity to expose, develop and carefully realign the holograms meant that the systems could not easily be updated to search for new objects. Filter updating is also needed if a system is to recognise different types of object or if the object under surveillance suffered an out-of-plane rotation. Even if the object suffers an in-plane rotation or a scale change then the Fourier plane filter needs to be changed if the image could not be de-rotated or re-scaled prior to the correlation process. Systems without the ability to update their filters will be referred to as fixed filter systems.

Even with their limitations, there were still tasks that could be performed by these fixed filter correlators. One of these was in the field of automatic photogrammetric stereo compilation [1.7] where, since the images being correlated formed a stereo pair, there was a guaranteed match between the object and reference images. Other proposals included word recognition [1.8], medical screening [1.9] [1.10], tracking and recognition of industrial components [1.11][1.12][1.13], autonomous navigation [1.14], cloud motion analysis and water pollution monitoring [1.15]. At the same time it was demonstrated that fixed filter correlators could be made very small, compact and robust [1.16] [1.17] which would be essential for operation in non-laboratory conditions.

The limitations imposed by fixed filters are self evident especially if one considers typical, military pattern recognition tasks. For example, if a ship or a tank is being engaged from the air then, it could be approached from any point on a hemi-spherical surface. However, a single MSF would only recognise the target from one particular

angle. In general there are two types of distortion that a recognition system needs to be made insensitive to:

1. In-plane distortions

- These include rotations about the line of sight and scale changes and shall be referred to here as 'type I distortions'.

2. Out-of-plane distortions

- These include any rotations with some component about an axis perpendicular to the line of sight and shall be called 'type II distortions'.

Some attempts to overcome distortion problems still used fixed filters. However, by using either multiplexing, averaging or coordinate transforms, a reduction was sought in the sensitivity of the system to the distortions. The multiplexing techniques included the addressing of a large number of filters in parallel, after optically replicating the input image [1.18][1.19][1.20], or even sequentially placing different filters into the Fourier plane [1.21]. The averaging techniques, which basically sought to make a single filter insensitive to distortions, included creating synthetic images from linear combinations of a set of training images (SDFs) [1.22] [1.23] [1.24]. These methods could be used to deal with both type I and type II distortions. Other techniques used the circular harmonic expansion of an image to produce filters (CHF's) that were insensitive to both in plane rotations and image contrast [1.25] [1.26] [1.27]. Filters were also designed that produced a constant signature as the filter was rotated in the Fourier plane and these could be made invariant to in plane distortions and image contrast [1.28] [1.29]. The most commonly used coordinate transform was the log-polar transform which could be used to overcome all type I distortions. By performing the log-polar transform on the FT of an image and correlating the resulting pattern, shift invariance could also be built in [1.30] [1.31] [1.32].

These were the main techniques for extending the operation of correlators using fixed filters. However, for all these techniques their ultimate limitation (even allowing for some of their practical limitations [1.33] [1.34] [1.35]) is that, in order to change the

object being searched for, the holographic filter or bank of filters has to be physically removed from the system and a new one inserted. This is a far cry from the total flexibility offered by digital systems. Ideally, to enable the optical correlator to access a large library of reference images, the holographic plate needs to be replaced by another medium. This should be capable of storing the FTs while possessing the ability to be erased and re-written *in situ* in the optical system.

Two methods of accomplishing this involve the use of electronically addressed spatial light modulators (SLMs: see later) in the Fourier plane of the correlation system. Both of these have been the subject of many published papers in recent years. The first is based on a Joint Transform Correlator (JTC) architecture [1.36]. Instead of allowing the reference and object images' transforms to fall onto a holographic plate, they are detected by a CCD camera. The resulting interference pattern is then relayed to a binary spatial light modulator [1.37]. However, this technique suffers from a chronic lack of space-bandwidth-product (SBWP) with image sizes limited to less than 64×64 pixels. This is due to the need to sample the interference pattern with the CCD detector [1.38]. The second method involves the use of a binary SLM to display phase only filters in the Fourier plane of the correlator [1.39][1.40]. Generally, using this technique, the amplitude information of the FT is discarded and the phase is quantised according to the dictates of the particular researcher. This phase pattern can then be reproduced by appropriately aligning the output polariser on a device such as a magneto-optic spatial light modulator [1.39]. This technique has a real potential advantage in that the reference functions can be stored as binary representations of FTs providing great savings in the memory requirements for their storage. Another great benefit is that the FT is sent directly to the SLM which removes the need for an interferometric system to record the FT of the reference image. This in turn relaxes the vibrational stability requirements of the system while allowing the use of cw lasers. Furthermore, because the FT is written to the SLM, there is no restriction on the use of complex valued inputs to form the reference templates. This opens up the possibility of using filters such as SDFs or CHF's in the updatable correlator [1.41] [1.42]. However, at the time of commencement of

this work, the SLM devices that were commercially available were limited to around 128×128 pixels and evaluation of a device purchased for another project had shown it to be of relatively poor quality. Two published reports have also highlighted problems with the current technology [1.43] [1.44]. The main objective of this research was to exploit the parallelism of the optical correlator in the processing of images with a large SBWP. It was therefore decided not to pursue this SLM approach but to note its potential for the future.

In order to utilise large SBWP images, it was decided to use an optically non-linear material to act as an updatable hologram in the Fourier plane. This material should, in the same way as a holographic plate, undergo physical changes as a result of the fringe patterns falling upon it. These physical changes may then be used to modulate a read out beam; thus allowing the stored hologram to be interrogated. The material chosen for the project was Bismuth Silicon Oxide (BSO) which has been used in the field of Fourier based optical image processing since the late 1970's. Its main advantage, for optical image processing, is its 'near photographic' sensitivity which allows reasonably fast hologram formation at moderate optical powers. Typical values are often quoted as a 10-40ms response time at $100\text{-}500\mu\text{Wcm}^{-2}$. Another advantage is that crystals of good optical quality can be obtained and the material possesses a high spatial resolution. From the point of view of the current work, an important factor was that good theoretical models had been developed for the hologram formation process. This allowed the definition of experiments for the measurement of those operating parameters pertinent to the correlator design.

Before presenting an outline history of the use of BSO in optical Fourier plane processors, it is useful to explain the difference between the two main correlator architectures that are employed. These are the Joint Transform Correlator (JTC) [1.36] and the Vander Lugt Correlator (VLC) [1.6]. The difference between these two architectures is illustrated in figure 1.2. In the case of the JTC the hologram is formed by placing the two input transparencies side by side in a collimated laser beam, in the front focal plane of a positive lens. The correlation plane is formed by reading out the hologram with a plane wave.

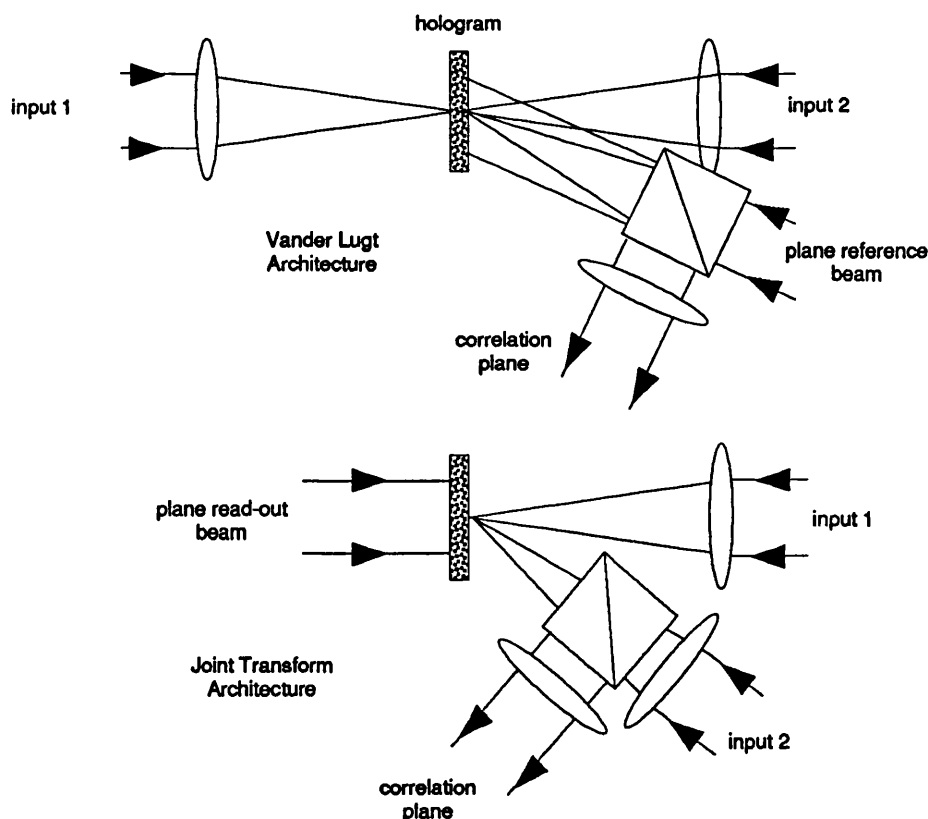


Figure 1.2: A comparison between the Vander Lugt Correlator (top diagram) and the Joint Transform Correlator (bottom diagram).

This read-out beam could be from a laser of different wavelength to that used for hologram writing. In the VLC however, the hologram is formed between two mutually coherent beams, one of which is the FT of one of the images while the other is a plane-wave reference. This hologram is then read out with the FT of the second image which again need not be coherent with the writing beams. However, in the case of the VLC, if a different read-out wavelength is used, steps must be taken to ensure that the scale of the object and reference FTs are matched.

1.1.3 A History of BSO-Based Correlators

A literature search was carried out, specifically targeted at the use of BSO in optical Fourier plane processors. These included systems for both correlation and other spatial filtering operations. The references are described below, in chronological order, to produce an historical overview of work in this field.

In 1978 Pepper et al [1.45] put forward the idea of using degenerate four-wave

mixing for the convolution and correlation of optical fields. This idea had grown out of original work on the use of four wave mixing for phase conjugation. In this paper it was demonstrated how manipulation of the equations for phase conjugation lead to convolution and correlation products. The proposed correlator used a JTC architecture. BSO was not mentioned and it is interesting to note that a theoretical example involving the use of CS_2 as the non-linear medium used input powers in the MW range. This paper also derived limits on the space bandwidth product of the system based on the original use of the Fresnel approximations in the optical FT. It pointed out the potential problems posed by the thickness of the material in terms of the need to preserve the form of the FT throughout the interaction region. The use of a single frequency source to overcome Bragg matching problems was also discussed.

This work, carried out at the California Institute of Technology, was followed up by White and Yariv, from the same institute, in 1980 [1.46]. This paper reported an experimental demonstration of the theoretical predictions using a JTC with photorefractive BSO at the Fourier plane ('suggested by the analogy between four-wave mixing and real time holography'). Although all three beams entering the crystal were mutually coherent, it was pointed out that the optical activity of the BSO reduced the interference between the plane read-out beam and the FT beams to negligible levels. The BSO crystal was 3mm thick and, to enhance the grating, a voltage of 5-7kV was applied to the crystal. The input laser used was a 1.6W Argon Ion (Ar^+) and response times of 30ms were reported (there being no indication of powers at the BSO however). Another practical point was the use of a polariser to isolate the desired, diffracted signal from other scattered light in the correlation plane. The correlations and convolutions were demonstrated using simple shapes and letters on photographic transparencies.

In 1981 a paper by Pichon and Huignard, of Thompson-CSF, was published [1.47]. This described a JTC but with wavelengths of 488nm and 633nm used for writing and reading the holograms respectively. A wavelength selective filter was added prior to the correlation plane to remove scattered light from the writing beams. This

was essential since the read out beam entered the BSO from the same side as write beams. The BSO was 2mm thick and the paper mentions the problems presented by Bragg matching requirements. The writing power was quoted as $200\mu W$ at the hologram. The system was found to have a response time of around 70ms (quoted as a time constant of $\sim 25\text{ms}$). Like Yariv's system it's operation was demonstrated with letters on photographic transparencies as inputs.

A general paper by White and Yariv in 1982 [1.48] covered areas including phase conjugation, image addition/subtraction and image plane edge enhancement. The section on correlation contained only the original 1980 results with the same JTC architecture. It also mentioned the potential problems presented by the interaction of the beam reading out the hologram with the hologram itself. A calculation was performed to determine the field of view constraints due to Bragg matching problems when one of the writing beam angles was changed.

An overview paper by Feinberg [1.49], from the university of Southern California, in 1983 discussed the use of photorefractives from a holographic rather than a four-wave mixing standpoint. The advantages of photorefractives were cited as their coupling efficiency, their stability and their erasability. Their disadvantages were listed as their slow speed at low intensities and the lack of understanding of the charge transport mechanism. The results used to illustrate the paper were again Yariv's of 1980. Also in 1983 Valley and Klein, of Hughes Research Laboratories, published a paper [1.50] in which Kukhatarev's band transport model was used. Figures of merit, relevant to optical data processing systems, were evaluated using the model and a comparison was made between BSO and BaTiO_3 . No reference to any correlation work at Hughes was mentioned however.

Two papers by Laycock and co-workers appeared in 1984 highlighting the interest of GEC's Hirst Research Centre in updatable optical correlators [1.51][1.52]. The papers described a JTC architecture with gratings being read and written at 633nm. The motivation for the use of the Helium-Neon (He-Ne) laser was the production of a small system with low power consumption. However, the performance must have suffered with the BSO being relatively insensitive at 633nm and the holographic

grating being equally affected by the read and write beams. Using a 2.7mm thick crystal with no applied field, a diffraction efficiency of 0.1% was reported, with a response time of 0.1 seconds. A half wave plate was used in the read beam to optimise its polarisation in order to maximise diffraction efficiency. To reduce the effects of the washing out of the grating by the read beam, this was chopped so as to be present for just a small fraction of the cycle time. Another significant step in this paper was the inclusion of a Hughes liquid crystal light valve in the system to give the potential for 'real time' input. The SLM was only used for the input of one of the images and was itself written to by imaging a transparency onto its back face. Also in 1984 GEC workers published a paper jointly with Hall et al of King's College, London [1.53]. This used the same JTC architecture but with no SLM and the He-Ne replaced with an Ar^+ laser. The paper dealt with the fidelity of the correlations formed and compared simulated and experimental results. In particular, the role of the relative beam intensities and the coupling of the read and write beams were examined. It was shown to be desirable to minimise the coupling between the read and write beams. This could be achieved by reducing their mutual coherence or by selecting an appropriate BSO axis. The controlling of spatial frequency response by the read beam intensity was discussed but it was not pointed out that the response would be much more easily controlled in a VLC architecture. For accurate correlation it was shown to be desirable to use a strong read beam intensity although this would obviously reduce the diffraction efficiency of the grating. The paper also noted that a fourth beam could be used to selectively erase portions of the grating to produce various degrees of edge enhancement. This was an idea that appeared later in a scheme designed specifically for edge enhancing images [1.54]. Finally, it pointed out that the thickness of the crystal should be small with respect to the overlap region of the writing beams.

In 1985 another correlator employing a SLM was reported by Thomson-CSF [1.55]. The architecture was again a JTC system with Ar^+ write and He-Ne read beams incident from the same side of the BSO. The input images were again static, being projected photographic transparencies. Apart from the SLM, the system con-

struction and performance were similar to that reported in their previous paper. The Bragg matching constraints were cited as a factor which would ultimately limit the number of pixels that could be processed in parallel. Also in 1985 Ochoa, Goodman and Hesselink - of Stanford University, California - demonstrated that variable degrees of edge enhancement could be produced by varying the strength of a plane reference beam in the recording of a FT hologram in BSO [1.56]. The particular application was the enhancement of defects in a periodic mask. The process relied on the fact that the diffraction efficiency of the BSO hologram was proportional to the reference to object beam ratio. Thus, by using a weak reference beam the weaker Fourier orders associated with the defects would be enhanced whilst the high intensity regions of the FT, which were due to the periodic mask structure, would be suppressed. However, no extension of this idea to the performance of spatial-frequency-weighted correlation was put forward in this paper.

In 1986, another overview of the use of photorefractive materials in optical data processing was published, this time by Fainman and Lee [1.57]. It covered areas such as matrix operations, statistical pattern recognition, associative memories and digital optical processing. The small section on correlation was again illustrated by the 1980 Yariv results but the more recent Thomson-CSF paper was also referenced. However, 1986 saw a number of papers published by workers from King's College, GEC and British Aerospace. A paper by Foote et al (King's college and BAe) [1.58] presented a review of the band transport model making a distinction between the intrinsic material parameters and the extrinsic system parameters. The predictions of the model were related to dynamic holographic optical processing applications including phase conjugation and correlation. The paper then went on to spell out the differences between processing schemes using the VLC and those using the JTC. In particular, it showed that the use of a VLC made it easier to control the problems created by the different growth times of the gratings corresponding to different Fourier frequencies. The paper pointed out that if it is desired to form a perfect auto-correlation then the JTC would give the best diffraction efficiency since fringe contrast would be maximised across the whole transform. However, if it was re-

quired to correlate an unknown scene against a library of references, then the VLC was useful since the reference beam could be used to maintain the fringe contrast at some required Fourier plane frequency. Also, since the hologram was read with the FT of the reference image, which merely needed to diffract from the already formed grating, the reference could be updated arbitrarily quickly. This last point was reiterated in another paper by Foote, Hall and Connors [1.59] which compared the growth and decay of the photorefractive grating with the cycle time achieved with a rapidly chopped beam diffracting from the hologram. While the JTC system would take tens of milliseconds to respond if either input image were changed, the VLC could operate with one image updated at arbitrarily fast rates.

A GEC paper by Cooper, Laycock et al now reported a VLC architecture. The advantages of the VLC discussed in the previous two papers were cited and spatial frequency weighted correlation was demonstrated. The correlator used a 633nm read beam which, it was pointed out, did not significantly perturb the BSO hologram. The hologram writing was accomplished using either a chopped 514nm beam or a 532nm pulsed laser source. The advantages of the pulsed system were the relaxation of stability requirements and an improvement in signal to noise ratio. This latter point was due to the absence of any intense writing beams during the hologram read out. This in turn meant that there was no light to be scattered into the correlation plane from defects in the BSO or other optics. The first experiment [1.60] simulated a pulsed laser by the use of an acousto-optic device to chop a cw Ar⁺ beam. The second [1.61] used a pulsed, frequency-doubled Nd:YAG source. The reported rise time for the BSO hologram was ~ 200 nanoseconds. All the inputs were again simple shapes on photographic transparencies. Two more papers, based on this system, appeared in 1987 [1.62] [1.63]. The two papers contained mostly the same material, beginning with a mathematical model of correlation by non-degenerate four-wave mixing in a VLC architecture. Coupled wave theory was used to examine, at a fairly basic level, the Bragg matching effects for both the reading and writing image transforms. Experimental measurements of the angular selectivity of holograms in BSO were presented. It was claimed that the selectivity was only half that predicted by theory

but no concrete explanation was put forward for this. Correlation results for the simple test shapes were shown, although any effects due to the Bragg matching constraints were not really evident.

Also in 1987, Vainos and Eason demonstrated multiple correlations performed simultaneously by multiplexed matched filters in a single BSO crystal [1.64]. The output correlation planes were spatially separated. In the example presented, two images, each consisting of four numerals, were correlated with a single numeral reference but it was claimed that the process could be extended to many more images.

Aldridge reported on correlation work carried out at British Aerospace [1.65]. Experimental results were presented demonstrating the dependence of diffraction efficiency on fringe modulation and of the response time on input intensity. Results for real time edge enhancement and weighted correlation using a VLC were also shown; again using photographic transparencies. It was noted that, if the reference beam was balanced to the higher spatial frequencies in the Fourier domain, then the system response time would increase because of their relatively low intensity.

The last paper from the search was published in 1988 by Coupland and Haliwell from Southampton university [1.66]. In this, BSO was used as the holographic medium for auto-correlating particle image velocimetry transparencies using a VLC. This paper was significant in the fact that the correlator was being applied to the solution of a real image processing problem. However, it should be noted that the use of a holographic medium for the performance of auto-correlation is, theoretically, unnecessary. This is because the auto-correlation function of an object is just the Fourier transform of its power spectrum. Thus, an optically addressed SLM could be used to convert the amplitude and phase of an optically produced FT into an intensity pattern which could then be optically re-transformed.

The overall picture to emerge from this search was that no new practical correlation systems were reported from the USA subsequent to the 1980 paper by Yariv. The main systems work in the UK had taken place at GEC between 1984 and 1987. However, this work then ended without any reported correlation system using real time input from an SLM. Furthermore, the reported results involved inputs that were

simple binary shapes. The work at Thomson-CSF had benefited greatly from their possession of SLMs, developed in-house, of sufficient quality for optical correlation. At the time of commencement of this research, Thomson were working on an ESPRIT funded program to develop a compact version of their JTC for industrial inspection tasks. However, it was thought that their SLMs were only capable of providing images with a SBWP of less than 300×300 pixels. Furthermore, the JTC architecture did not easily lend itself to the production of weighted correlation functions. Thus, the construction of a BSO based optical correlator for use, with real time inputs, in a high resolution industrial inspection task was considered to be a potentially large step forward in this technology area. The use of weighted correlations to provide varying degrees of discrimination would also aid the interpretation of the results.

1.1.4 Spatial Light Modulators and Input Images

Although the use of BSO in the Fourier plane allows the matched filters to be updated in real time, there is still one crucial component needed for a real time system. This is a spatial light modulator (SLM), the function of which is to convert input images into an intensity modulation of a collimated laser beam. The original images may have been imaged directly onto a light sensitive face of the SLM, acquired by a CCD camera or generated by a computer. It is not intended to go into detailed descriptions of these devices here. For general information on spatial light modulators the reader is referred to [1.67][1.68] and [1.69].

One topic that will be covered in this thesis is the effect on system performance of the optical quality and the fidelity of these devices. For this study it should be noted that the spatial light modulator may not exactly reproduce the intensities of the original scene. Localised changes in the phase of the beam may result from the mechanisms used to modulate the beam's intensity. It may also, like any of the other optical components in the system, be subject to defects such as a lack of overall flatness or surface roughness. The device intended for use in this research was of a type that behaves like an updatable photographic transparency. The SLM is optically addressed by an intensity pattern which falls onto a photoconductive layer. The local

changes in conductivity, thus caused, are used to spatially modulate the strength of an electric field applied across a liquid crystal layer. The field modulation in turn produces a two dimensional pattern in the alignment angles of the liquid crystal molecules and this modulates the polarisation state of a collimated laser beam. A polariser placed after the SLM can then be used to create the intensity image. This is as opposed to devices which comprise a matrix of individual liquid crystal cells addressed by an electrode grid to produce a pixellated image. In order to relay images, in real time, from a TV system to the SLM, the images are displayed on a miniature cathode ray tube (CRT). The CRT's output can then be imaged onto the photoconductive layer of the SLM either by conventional optics or by a fibre optic faceplate. The choice of the CRT determines the spatial frequency of the input images; since the resolution of the CRT was expected to be somewhat lower than that of the SLM.

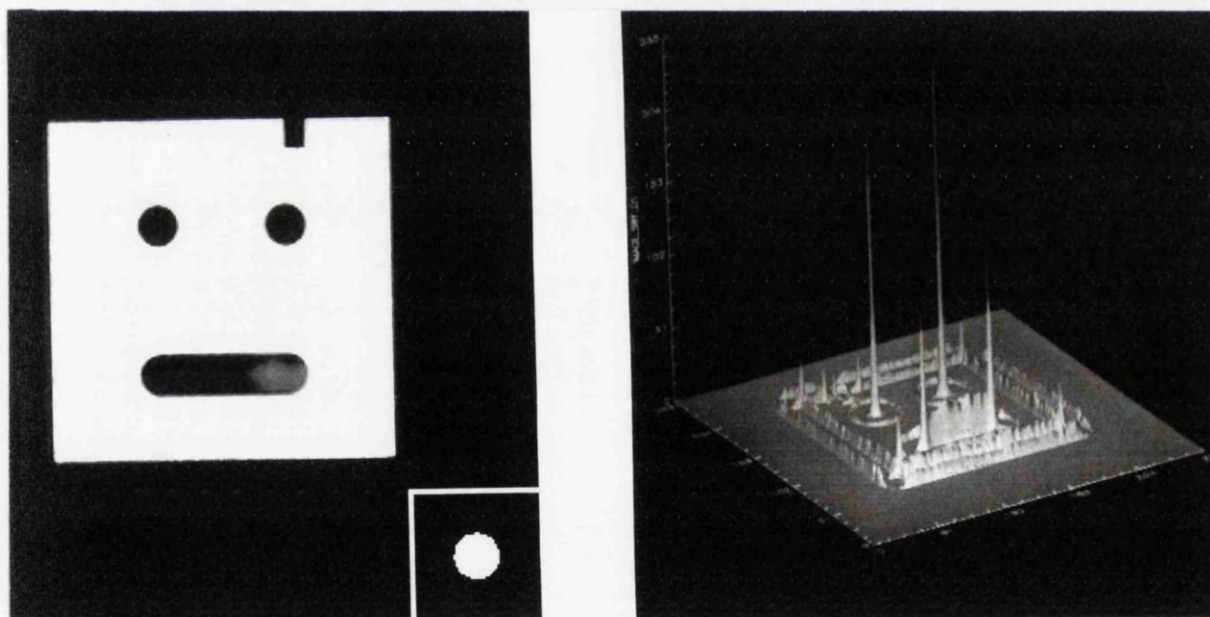


Figure 1.3: Images of machined component (left) with simple computer generated hole template inset and their computed cross-correlation (right).

In any discussion of the performance of an optical correlator, the conclusions will nearly always relate closely to a specific image processing task. As mentioned at

the beginning of this chapter, the task considered for this research was a simulated industrial inspection process. A generic, machined-metal component was compared to a computer-generated model. As one part of this comparison, the various faces of the machined component were searched for a sub-feature from a library of images. An example of the imagery used is shown on the left of figure 1.3, with a simple, computer-generated hole template inset in the top left corner. On the right of the figure is a computer generated cross-correlation, formed from edge enhanced versions of the two input images.

The crucial feature of the correlation plane is the ratio of the peak intensity to the background intensities in the rest of the correlation plane. The magnitude of the peak should be such that simple thresholding can highlight its presence whilst eliminating the rest of the correlation plane structure. It should be noted that this background structure stems from two sources: the first, is noise which may be introduced either in the input images or in the processing; the second, is the structure due to the cross-correlation of the reference image with the whole input shape. Noise can often be reduced by careful system design. However, the cross-correlation structure, which dominates the example of figure 1.3, can never be reduced below its minimum level. It will, however, take different forms depending on the nature of the edge enhancement applied to the images.

It was decided that the optical correlator should be capable of processing grey-scale images of up to 512×512 pixels. It was also decided that the correlator should have the capability of cross-correlating two 512×512 pixel images; even though, in the example task, the reference template was small compared to the input image. Since the images were being acquired by a CCD camera it was also decided that the target cycle time for the system would be as close as possible to 40ms which is equal to a TV frame time.

1.2 The Photorefractive Effect and the Band Transport Model

As discussed in the previous section, BSO was selected as the medium in which updatable, holographic matched filters would be formed in the Fourier plane of the correlator. The mechanism by which BSO is able to form temporary or fixed [1.70] holograms is the photorefractive effect. An advantage gained in the selecting BSO was that the physical process of hologram formation had been extensively modelled by a number of researchers. This allowed the work to concentrate on the correlator design. One of the most complete models of this process is the 'Band Transport Model' (BTM) [1.71]; although a 'Charge Hopping Model' [1.72] has also been put forward. The actual mathematical formulation of the BTM is not included in this thesis but may be found elsewhere [1.71][1.73][1.74]. It should be noted that [1.73] is, itself, a review paper and contains a further eighty two references. Although the mathematical details are not included, it is thought that a description of the photorefractive mechanism may be of use to any reader of this dissertation.

It is assumed that defects exist within the BSO crystal, e.g. lattice vacancies or atoms of some foreign substance. These defects may be classified into three groups:

1. *acceptor sites* which may receive and trap electrons, becoming negatively charged in the process.
2. *donor sites* which may lose electrons into the conduction band of the crystal.
3. *traps* which are those donor sites that have become ionised.

An electron may be lost from a donor site as a result of either photo-excitation or by thermal excitation. The ionised donor site (trap) may accept and trap electrons although, of course, these could be re-excited again at some later time. The presence of the acceptor sites is important because they can trap and hold thermally excited electrons. It is usually assumed that every available acceptor site in the crystal has been filled by an electron which in turn has come from a donor site. The significance

of this is that, even in the dark there will be both donors and traps present in the crystal.

Consider a situation in which one part of the crystal is illuminated while its neighbouring regions are not. In the illuminated region electrons will be excited into the conduction band in which they are free to move about the crystal. Since there will be a relatively high local concentration of conduction band (free) electrons in the illuminated region, there will be a net diffusion into the dark regions. Electrons in the conduction band may also be encouraged to move by the application of an electric field to the crystal. These electrons will eventually be captured in a trap site; however electrons captured in traps in the illuminated region are liable to be re-excited again. On the other hand, electrons which become trapped in the dark regions are less likely to be moved on if the rate of thermal excitation is not too large. Thus, electrons are gradually moved out of the illuminated and into the dark regions. This displacement of charge causes an internal electric field to build up between the dark and the light regions. This field itself will start to inhibit the exodus of electrons from the illuminated regions. Eventually a stalemate will be reached when the forces of the internal field exactly counter the diffusive 'force' and the force of any externally applied field.

At this stage there exists, within the non-uniformly illuminated crystal, a non-uniform electric field. This field will not, in itself, affect light passing through the crystal. However, BSO exhibits the linear electro-optic effect and its refractive index will be altered wherever there is an electric field present. Any non-uniform refractive index distribution within the crystal *will* affect incident light. Since, there will be a limit to the size of the electric fields that can be sustained, it is important to note that the refractive index change is linearly proportional to the magnitude of the electric field. The material parameter that links this refractive index change to the magnitude of the electric field is the effective electro-optic coefficient, r_{eff} . Things are not quite straight forward however, because r_{eff} is actually different depending on which orientation of the crystal is being considered. In order to describe the way in which the electric fields, and hence the changes in refractive index, build up, the

BTM may be used.

The starting point of this model is four equations. The first equation concerns the rate of change of local free electron density and it takes into account the formation and absorption of the free electrons and their ability to drift into and out of a local volume. It includes a term for the probability of electrons being excited which depends on several factors including incident intensity, local donor concentration and temperature. Another material parameter included in the rate equation is the photo-excitation constant. This is wavelength dependent, being relatively high for green light but many times lower in the red part of the spectrum. There is also a term for the probability of re-trapping which includes the local concentration of traps and also the density of free electrons. Finally there is the term describing electron drift. This includes such parameters as the diffusion coefficient, the difference in electron concentration between the region of interest and its neighbours, the electron mobility and the local electric field.

The second equation concerns the rate of change of local trap site density. It contains the same rate terms as the equation for electron generation, which implies the assumption that all free electrons present come from and are re-absorbed by donor atoms. There is however no drift term since the donor and trap sites are immobile within the crystal.

The third and fourth equations are Gauss's law and a quasi static approximation of one of Maxwell's equations. Gauss's law relates the derivative of the electric field, with respect to the spatial coordinates within the crystal, to the local net unbalanced-charge. The fourth equation fixes the direction of the electric field.

These equations illustrate an important fact, namely, that the parameters that will ultimately govern the performance of the BSO as an updatable holographic medium fall into two groups. Firstly, there are the parameters such as the diffusion constant and the donor concentration which are internal to the crystal itself. Unless a large number of different crystals are purchased and characterised, these will be unalterable 'facts of life' in the construction of the correlator. Secondly, there are parameters such as: the temperature, the incident intensity and the applied electric

field. These parameters may be adjusted when the correlator is set up in order to meet, within practical limits, the desired operating characteristics.

The problem with the equations of the BTM is that they are all nonlinearly coupled and so, in general, difficult to solve. It is therefore necessary to consider some simple situations for which approximations may be used to make solution of the equations possible. For hologram formation, the case to consider is that for cosinusoidal illumination. In this case, it is assumed that the incident light is a uniform grating with fringes in one direction only. As electrons move from the bright parts of the fringe pattern to the dark parts a 'grating like' electric field distribution will result which will in turn lead to a grating like refractive index modulation which is basically a hologram. This is illustrated in figure 1.4:

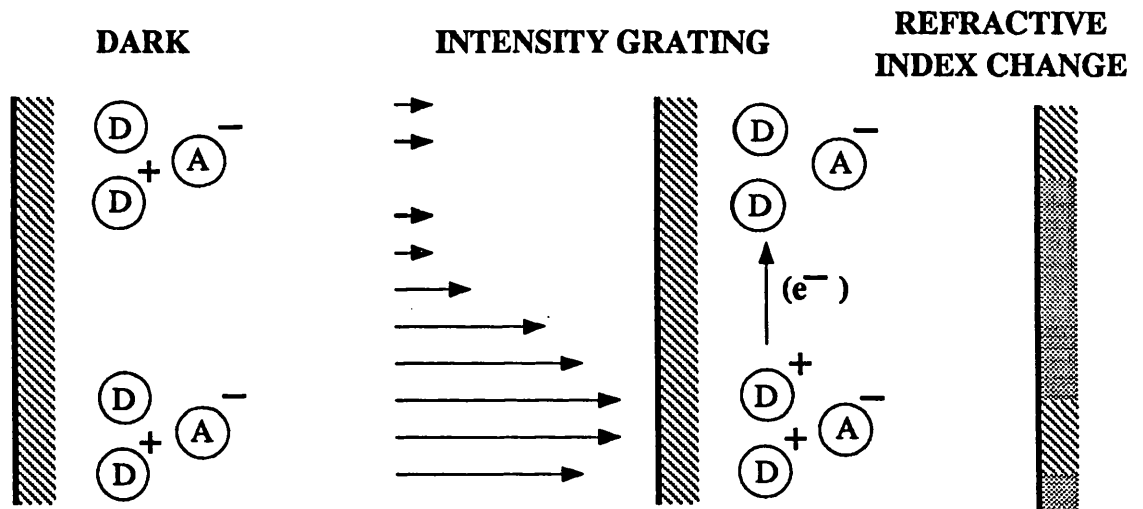


Figure 1.4: Grating formed by charge separation caused by cosinusoidal illumination.

The mathematical derivation of solutions to the equations of the BTM, including solutions for cosinusoidal illumination is found in [1.73]. The results that will be used in subsequent chapters are the expressions for the magnitude of the space charge field and its temporal behaviour.

1.3 The Objectives and Scope of this Research

As mentioned at the beginning of the chapter, this research was undertaken within the wider framework of a three year research program, part funded by the EEC. The research presented in this thesis concentrates on a thorough design and evaluation of a BSO based optical correlator. The starting points for the design were the band transport model and the results from coupled wave theory which were used to predict the effects of the thick BSO hologram in the Fourier plane. The work addressed every relevant practical consideration and lead to a design ready to be constructed on a laboratory optical bench.

The key system parameters were:

- the cycle time;
- the images' space bandwidth product;
- the fidelity of the correlation;
- and the light utilisation.

The variables used to achieve the desired parameters were:

- the laser powers and wavelengths;
- the basic correlator geometry;
- the strength of the external field applied to the BSO;
- the f-numbers of the lens systems;
- the holographic fringe spacing;
- and the polarisations of the write and read beams.

Chapter 2 of this thesis is concerned with the problems presented by the presence of a thick hologram in the Fourier plane of the correlator. It uses the results of Kogelnik's coupled wave theory [1.75] to predict both the maximum diffraction

efficiency and the variations in diffracted power that may be expected. The expected refractive index modulation is calculated, based on measurements made on a specific BSO crystal. Expressions are derived for the beam angles in the Fourier plane. The problems associated with the depth-of-field of the Fourier transform are also discussed.

Chapter 3 considers other factors pertinent to the correlator architecture which were studied as the design work progressed. These included problems presented by the location of the input images, which are discussed for the first time with respect to optical correlators employing thick holograms. It is also shown how computer simulation can be used to determine the optical powers present in the various planes of the system. This is used in the design process to specify laser powers. Experimental results are presented which allow the prediction of the BSO's response time for any given laser power, applied electric field and fringe spacing. The predictions of the BTM and the coupled wave theory also lead to requirements for optimising holographic fringe spacing. These are compared to the requirements that are imposed by the need to separate the correlation plane from the undiffracted read-beam. One effect, not previously considered in the design of a BSO-based correlator, is that of holographic fringes becoming tilted with respect to the applied field. This has implications for the image being written into the BSO and the problem is discussed and an experimental evaluation carried out. The way in which FTs are stored in BSO is compared to their storage in conventional photographic media. A computer simulation is presented which leads to some conclusions regarding the effect of the holographic storage on spatial filtering of the FTs.

Once produced, the final design was to have been experimentally evaluated with respect to correlations performed as part of an industrial inspection task. However, a major problem was encountered with a failure to produce the spatial light modulators for the correlator. The work on the optical system had to be abandoned in order to concentrate on other areas within the framework of the overall project. The effort that was to have been put into the construction and evaluation of the optical system was switched to a computer simulation of spatial light modulator distortions.

This work provided simulation methods which could be used to assess the requirements for SLM quality for the performance of given tasks. It was also concerned with demonstrating a suitable experimental method to assess the effects of distortions.

Chapter 4 discusses three areas, with implications for correlators employing optically addressed SLMs. These are: device resolution; large and small scale phase distortions; and grey-level distortions. In order to carry out a simulation of the effects of these distortions on correlation, the 2^N factorial method was proposed as a suitable tool. The advantages of this technique are discussed, along with its limitations as regards a computer simulation. An example experiment was performed to assess SLM distortions relative to the location of a sub-feature – namely a drilled hole – in the image of a machined component. In this experiment, three distortion parameters were modelled and their direct effects on the correlation task were assessed as well as the interactions between them.

Chapter 5 relates the results derived in the preceding chapters to the actual system parameters that will apply to the optical correlator. An actual system design is produced, based on the use of 'off-the-shelf' optical components. The equations governing the correlator's performance, are expressed in terms of a few key system parameters. This allowed some predictions to be made about the potential for processing larger SBWP images in a BSO correlator. The reduction in the number of parameters is based on the assumption that the lens' focal length is selected to match the highest Fourier plane spatial frequency to the radius of the BSO crystal.

References: Chapter 1

- 1.1 "Digital Image Processing", Pratt, W.K., John Wiley and Sons, 1987.
- 1.2 "Digital Image Processing and Computer Vision", Schalkoff, R.J., John Wiley and Sons, 1989.
- 1.3 "Digital Image Processing (Second Edition)", Gonzalez, R.C. and Wintz, P., Addison Wesley, 1987.
- 1.4 "The Fourier Transform and its Applications (Second Edition)", Bracewell, R.N., McGraw-Hill Kogakusha, 1978.
- 1.5 "Introduction to Fourier Optics", Goodman, J.W., McGraw-Hill, 1968.
- 1.6 Vander Lugt, A., "Signal Detection by Complex Spatial Filtering", IEEE Trans., IT-10 p139 (1964).
- 1.7 "Optical Data Processing", Casasent, D.(Ed.), Topics in Applied Physics Vol.23 chapt.5, Springer-Verlag, 1978.
- 1.8 Casasent, D., Caimi, F., Hinds, J.E., "Optical Word Recognition: Case Study in Coherent Optical Pattern Recognition", Opt.Eng., 19 p716 (1980).
- 1.9 "Optical Data Processing", Casasent, D.(Ed.), Topics in Applied Physics Vol.23 p235, Springer-Verlag, 1978.
- 1.10 Parker, S.C.J., "The use of Optical Matched Filtering for the Detection of Cervical Cancer", J.Phys.D: Appl.Phys., 21 pS174 (1988).
- 1.11 Gibson, D.G., Aldridge, N., Brown, M.B., Levenston, A.G., White, H.J., "Automatic Recognition and Tracking of Targets from Visible and Thermal Imagery, Using Optical Processing", Proc.SPIE, 492 ECOOSA '84 p.165 (1984).
- 1.12 Gara, D.A., "Real Time Optical Correlation of 3-D Scenes", Appl.Opt., 16 p.149 (1977).

- 1.13 Gara, D.A., "Real Time Tracking of Moving Objects by Optical Correlation", Appl.Opt., **18** p.172 (1979).
- 1.14 Casasent, D., Luu, T.K., "Photo-DKDP Light Valve in Optical Data Processing", Appl.Opt., **18** p.3307 (1979).
- 1.15 "Optical Information Processing", Lee, S.H.(Ed.), Topics in Applied Physics Vol.48 p204-207, Springer-Verlag, 1981.
- 1.16 Duthie, J.G., Upatnieks, J., "Compact Real-Time Coherent Optical Correlators", Opt.Eng., **23** p.007 (1984).
- 1.17 Gregory, D.A., Kirsch, J.C., "Compact Optical Correlators", Proc.SPIE, **960** Real Time Signal Processing for Industrial Applications p.66.
- 1.18 Leib, K.G., Bondurant, R.A., Hsiao, S., Wohlers, R., Herold, R., "Aerial Reconnaissance Film Screening Using Optical Matched Filter Image Correlator Technology", Appl.Opt., **17** p.2892 (1978).
- 1.19 Chao, T-H., Liu, H-K., "Real Time Optical Holographic Tracking of Multiple Objects", Appl.Opt., **28** p.226 (1989).
- 1.20 Cawte, P.S., Cooper, I.R., Gibbons, G.G., Webster, S.C., Dubois, F., Buchel, M., Brodin, J.M., Schnell, J.Ph., Loiseaux, B., "An Optical Image Correlator for Robotic Applications", Proc.SPIE, **1134** Optical Pattern Recognition II p.196 (1989).
- 1.21 Psaltis, D., Neifield, M.A., Yamamura, A., "Image Correlators Using Optical Memory Disks", Opt.Lett., **14** p.429 (1989).
- 1.22 Casasent, D., "Unified Synthetic Discrimination Function Computational Formulation", Appl.Opt., **23** p.1620 (1984).
- 1.23 Casasent, D., Chang, W.T., "Correlation Synthetic Discrimination Functions", Appl.Opt., **25** p.2343 (1986).

- 1.24 Tian, Q., Fainman, Y., Gu, Z.H., Lee, S.H., "Comparison of Statistical Pattern Recognition Algorithms for Hybrid Processing. I. Linear Mapping Algorithms", J.O.S.A.(A), **5** p.1655 (1988).
- 1.25 Hsu, Y-N., Arsenault, H.H., April, G., "Rotation Invariant Digital Pattern Recognition Using Circular Harmonic Expansion", Appl.Opt., **21** p.4012 (1982).
- 1.26 Arsenault, H.H., Delisle, C., "Contrast Invariant Pattern Recognition Using Circular Harmonic Components", Appl.Opt., **24** p.2072 (1985).
- 1.27 Sheng, Y., Arsenault, H.H., "Object Detection from a Real Scene using the Correlation Peak Coordinates of Multiple Circular Harmonic Filters", Appl.Opt., **28** p.245 (1989).
- 1.28 Sweeny, D.W., Ochoa, E., Schils, G.F., "Experimental use of Iteratively Designed Rotation Invariant Correlation Filters", Appl.Opt., **26** p.3458 (1987).
- 1.29 Mahalanobis, A., Vijaya Kumar, B.V.K., Casasent, D., "Spatial Temporal Correlation Filter for In Plane Distortion Invariance", Appl.Opt., **25** p.4466 (1986).
- 1.30 Psaltis, D., Casasent, D., "Deformation Invariant Optical Processors Using Coordinate Transforms", Appl.Opt., **16** p.2288 (1977).
- 1.31 Casasent, D., Psaltis, D., "Position Rotation and Scale Invariant Optical Correlation", Appl.Opt., **15** p.1795 (1976).
- 1.32 Saito, Y., Komatsu, S., Hitoshi, O., "Scale and Rotation Real Time Optical Correlator Using Computer Generated Hologram", Opt.Comm., **47** p.8 (1983).
- 1.33 Mengel, F., "Resolution Limits of Holographic Mapping Filters for Geometric Coordinate Transformations", Appl.Opt., **27** p.339 (1988).
- 1.34 Riggins, J., Butler, S., "Simulation of Synthetic Discriminant Function Optical Implementation", Opt.Eng., **23** p.721 (1984).

- 1.35 Reid, M.B., Ma, P.W., Downie, J.D., Ochoa, E., "Experimental Verification of Modified Synthetic Discriminant Function Filters for Rotation Invariance", *Appl.Opt.*, **29** p.1209 (1990).
- 1.36 Weaver, C.S., Goodman, J.W., "A Technique for Optically Convolver Two Functions", *Appl.Opt.*, **5** p.1248 (1966).
- 1.37 Javidi, B., "Joint Transform Image Correlation Using a Binary SLM at the Fourier Plane", *Appl.Opt.*, **27** p.663 (1988).
- 1.38 Young, R., "Resolution Limitations Associated with the use of a CCD Array Detector in the Frequency Plane of a Joint Transform Correlator", HII (BRITE-EURAM P2202) project report, HII/TR/GLU/R901208.
- 1.39 Psaltis, D., Paek, E.G., Santosh, S.V., "Optical Image Correlation with a binary Spatial Light Modulator", *Opt.Eng.*, **23** p.698 (1984).
- 1.40 Horner, J.L., Leger, J.R., "Pattern Recognition with Binary Phase Only Filters", *Appl.Opt.*, **24** p.609 (1985).
- 1.41 Jared, D.A., Ennis, D.J., "Inclusion of Filter Modulation in Synthetic Discriminant Function Construction", *Appl.Opt.*, **28** p.232 (1989).
- 1.42 Leclerc, L., Sheng, Y., Arsenault, H.H., "Rotation Invariant Phase-Only and Binary Phase-Only Correlation", *Appl.Opt.* **28** p.1251 (1989).
- 1.43 Davis, J.A., Gamieli, J., Bach, G.W., "Optical Transmission and Contrast Ratio Studies of the Magneto optic Spatial Light Modulator", *Appl.Opt.*, **27** p.5194 (1988).
- 1.44 Hartman, N.F., Gaylord, T.K., "Coherent Optical Characterisation of Magneto optic Spatial Light Modulators", *Appl.Opt.*, **29** p.4372 (1990).
- 1.45 Pepper, D.M., Yeung, J.A., Fekete, D., Yariv, A., "Spatial Convolution and Correlation of Optical Fields via Degenerate Four-Wave Mixing", *Opt.Lett.*, **3** p.7 (1978).

- 1.46 White, J.O., Yariv, A., "Real Time Image Processing via Four-Wave Mixing in a Photorefractive Medium", *Appl.Phys.Lett*, **37** p.5 (1980).
- 1.47 Pichon, L., Huignard, J.P., "Dynamic Joint Transform Correlator by Bragg Diffraction in Photorefractive $\text{Bi}_{12}\text{SiO}_{20}$ Crystals", *Opt.Comm.*, **36** p.277 (1981).
- 1.48 White, J.O., Yariv, A., "Spatial Information Processing and Distortion Correction via Four-Wave Mixing", *Opt.Eng.*, **21** p.224 (1982).
- 1.49 Feinberg, J., "Optical Image Processing Using Four-Wave Mixing in Photorefractive Materials", *Proc.SPIE*, **388** Advances in Optical Information Processing p.106 (1983).
- 1.50 Valley, G.C., Klein, M.B., "Optimal Properties of Photorefractive Materials for Optical Data Processing", *Opt.Eng.*, **22** p.704 (1983).
- 1.51 Petts, C.R., McCall, M.W., Laycock, L.C., "Optical Correlation in $\text{Bi}_{12}\text{SO}_{20}$ at 632.8nm", *Electron.Lett.*, **20** p.32 (1984).
- 1.52 Laycock, L.C., McCall, M.W., Petts, C.R., "A Compact Real-Time Optical Processing System", *GEC Journal of Research*, **2** p.83 (1984).
- 1.53 Connors, L., Foote, P., Hall, T.J., Jaura, R., Laycock, L.C., McCall, M.W., Petts, C.R., "Fidelity of Real-Time Correlation by Four-Wave Mixing", *Proc.SPIE*, **492** ECOOSA '84 p.361 (1984).
- 1.54 Vainos, N.A., Eason, R.W., "Real Time Edge Enhancement by Active Spatial Filtering via Five Wave Mixing in Photorefractive BSO", *Opt.Comm*, **59** p.167 (1986).
- 1.55 Loiseaux, B., Illiaquer, G., Huignard, J.P., "Dynamic Optical Cross Correlator Using a Liquid Crystal Light Valve and a Bismuth Silicon Oxide Crystal in the Fourier Plane", *Opt.Eng.*, **24** p.144 (1985).
- 1.56 Ochoa, E., Goodman, J.W., Hesselink, L., "Real Time Enhancement of Defects in a Periodic Mask Using Photorefractive $\text{Bi}_{12}\text{SiO}_{20}$ ", *Opt.Lett.*, **10** p.430 (1985).

- 1.57 Fainman, Y., Lee, S.H., "Application of Photorefractive Crystals to Optical Signal Processing", Proc.SPIE, **634** Optical and Hybrid Computing p.380 (1986).
- 1.58 Foote, P.D., Hall, T.J., Aldridge, N.B., Levenston, A.G., "Photorefractive Materials and their Applications in Optical Image Processing", Proc.IEE, **133** (Pt.J) p.83 (1986).
- 1.59 Foote, P.D., Hall, T.J., Connors, L.M., "High Speed Two Input Real Time Optical Correlation Using Photorefractive BSO", Opt.Las.Tech., Feb.'86 p.39 (1986).
- 1.60 Cooper, I.R., Laycock, L.C., McCall, M.W., Nicholson, M.G., Petts, C.R., Powell, M.A., "A Dynamic Frequency Plane Image Correlator", Proc.IEE, **133** (Pt.J.) p.70 (1986).
- 1.61 Nicholson, M.G., Gibbons, G.G., Laycock, L.C., Petts, C.R., "Image Correlation via Pulsed Dynamic Holography", Electron.Lett., **22** p.1200 (1986).
- 1.62 Nicholson, M.G., Cooper, I.R., Gibbons, G.G., Petts, C.R., "Optimisation of an Updatable Optical Image Correlator", Opt.Eng., **26** p.445 (1987).
- 1.63 Nicholson, M.G., Cooper, I.R., McCall, M.W., Petts, C.R., "Simple Computational Model of Image Correlation by Four-Wave Mixing in Photorefractive Media", Appl.Opt., **26** p.278 (1987)
- 1.64 Vainos, N.A., Eason, R.W., "Spatially Multiplexed Phase Conjugate Imaging and Processing in Photorefractive BSO", Opt.Comm., **62** p.311 (1987).
- 1.65 Aldridge, N.B., "Optical Image Processing in GaAs and BSO", Proc.SPIE, **812** Progress in Holography p.36 (1987).
- 1.66 Coupland, J.M., Halliwell, N.A., "Particle Image Velocimetry: Rapid Transparency Analysis Using Optical Correlation", Appl.Opt., **27** p.1919, (1988).
- 1.67 "Optical Information Processing", Lee, S.H.(Ed.), Topics in Applied Physics Vol.48, Chapter 4, Springer-Verlag, 1981.

- 1.68 Fisher, A., "Spatial Light Modulators: Functional Capabilities, Applications and Devices", Int.Jour.of Optoelectronics, **5** p.125 (1990).
- 1.69 "Spatial Light Modulators and Applications", 1990 Technical Digest Series **14**, September 10-12 Nevada, Optical Society of America (1990).
- 1.70 Herriau, J.P., Huignard, J.P., "Hologram Fixing Process at Room Temperature in Photorefractive $\text{Bi}_{12}\text{SiO}_{20}$ ", Appl.Phys.Lett., **49** p.1140 (1979).
- 1.71 Kukhtarev, N., Markov, V.B., Odulov, S.G., Vinetskii, V.L., Ferroelectrics , **22** p.949 (1979).
- 1.72 Feinberg, J., Heiman, D., Tanguay, A.R., Hellwarth, R.W., Jour.Appl.Phys., **51** p.1297 (1980).
- 1.73 Hall, T.J., Jaura,R., Connors, L.M., Foote, P.D., "The Photorefractive Effect - A Review", Progress in Quantum Electronics, **10** p.77 (1985).
- 1.74 "Optically Induced Anisotropic Light Diffraction in Photorefractive Crystals", Foote, P.D., Ph.D. Thesis (University of London), (1987).
- 1.75 Kogelnik, H., "Coupled Wave Theory for Thick Hologram Gratings", Bell Syst.Tech.Jour., **48** p.2909 (1969).

Chapter 2

Thick Holograms in an Optical Correlator

Contents of Chapter 2

2.1	Maximum Diffraction Efficiency	50
2.1.1	Refractive Index Modulation of BSO	52
2.2	Variations in Diffraction Efficiency	56
2.2.1	Changes in Read-Beam Angle	57
2.2.2	Changes in Write-Beam Angle	60
2.2.3	Beam Angles Produced in the Correlator	66
2.3	Validity of the Use of Coupled Wave Theory	68
2.4	Depth of Field in the Fourier Plane	75
2.5	Conclusions	76

Abbreviations/symbols used in Chapter 2

BSO	Bismuth Silicon Oxide (Chapt. 1).
BTM	Band Transport Model (Chapt. 1).
CRT	Cathode Ray Tube (Chapt. 1).
MSF	Matched Spatial Filter (Chapt. 1).
SLM	Spatial Light Modulator (Chapt. 1).
B	Bragg angle for hologram.
ΔB	Difference between read-beam and Bragg angle.
d	Hologram thickness.
E_{SC}	Total electric field across BSO.
E_A	Electric field applied to BSO.
f	Lens' focal length.
n	Refractive index (of BSO).
Δn	Refractive index modulation (of hologram).
r_{eff}	Effective electro-optic coefficient of BSO.
u	Spatial frequency.
Q	Hologram thickness parameter.
Δz	Represents the longitudinal extent of the Fourier transform.
α	Half-angle between holographic writing beams, inside holographic medium.
$\Delta\alpha$	Difference between actual α and that required to produce fringe spacing Λ .

η	Diffraction efficiency.
Λ	Hologram's fringe spacing.
Λ_C	Critical fringe spacing for BSO.
λ	Free-space wavelength.
$\lambda_r: \lambda_w$	Subscripts to denote reading and writing of hologram.
ϕ	Holographic grating vector angle.

Unlike the typical matched spatial filters (MSFs), recorded as holograms on photographic emulsions, the MSFs recorded in the BSO crystal may display significant angular selectivity. This is because the physical thickness of the BSO may be of the order of millimetres as opposed to photographic emulsions which have thicknesses measured in microns. In order to predict the sort of effects that will be encountered in the correlator, a theory describing the diffraction of light from thick holograms is required. For this work, the coupled wave theory of Kogelnik [1.75] is chosen as it provides a relatively simple picture of the diffraction process. This theory results in an equation for the diffraction efficiency as opposed to other methods such as those of Bruckhardt [2.1] and Kaspar [2.2] which require numerical solutions. The coupled wave theory provided results which at least highlighted the main practical problems and quantified their relative severity. The first predictions in this section concern the maximum diffraction efficiency expected from the BSO. The effects of the hologram's angular selectivity are then assessed with respect to the angular spread of the reading and writing beams. Finally, the validity of using coupled wave theory for this work will be discussed and some experimental results presented.

2.1 Maximum Diffraction Efficiency

The following results for the diffraction efficiency of a lossless, thick transmission hologram of arbitrary fringe slant are presented in [1.75]. The results are derived from coupled wave theory and the definitions are with respect to the geometry of figure 2.1. For phase only holograms, it is assumed in the derivation that the average refractive index is large with respect to the refractive index modulation. Note that all the angles are defined *within the medium of refractive index n* and the grating boundaries do not represent a transition between regions of different refractive indices. The angles are measured with respect to the z axis which is defined as the normal to the boundaries of the grating region. The wavelengths are all measured in free space, i.e. *outside the medium*.

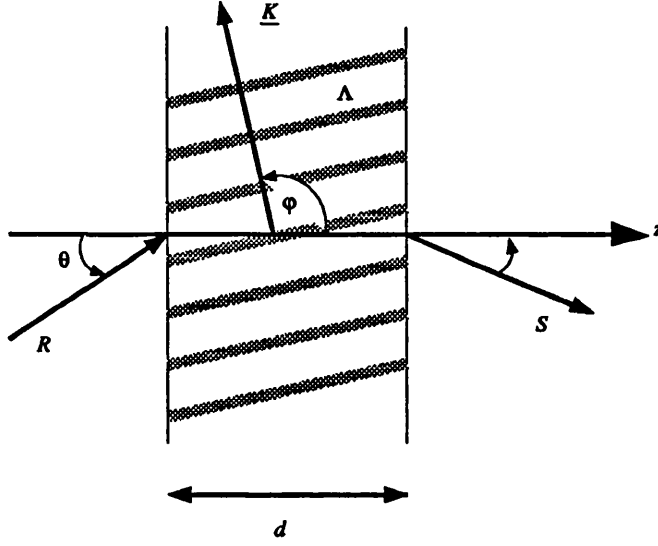


Figure 2.1: Geometry used to describe the diffraction from the volume hologram.

The diffraction efficiency, η , defined as the power in the diffracted wave produced by an incident wave of unit amplitude, is:

$$\eta = \frac{c_S}{c_R} S S^* = \frac{\sin^2 \sqrt{v^2 + \xi^2}}{1 + \frac{\xi^2}{v^2}} \quad (2.1)$$

where:

$$\begin{aligned} c_R &= \cos \theta & c_S &= \cos \theta - (\lambda_r \cos \phi / \Lambda n) \\ v &= \pi \Delta n d / (\lambda_r \sqrt{c_R c_S}) & \xi &= \Delta B \pi d \sin(\phi - B) / (\Lambda c_S) \end{aligned} \quad (2.2)$$

and B is the angle of incidence that satisfies the Bragg condition for maximum transfer from the incident wave to the diffracted wave. The angle ΔB is the difference between the incident angle and the Bragg angle, B , or the Bragg miss-match. The Bragg angle, B , can be obtained from:

$$\cos(\phi - B) = \frac{\lambda_r}{2\Lambda n} \quad (2.3)$$

When a wave is incident on the grating at the Bragg angle, $\Delta B = 0$ and the amplitude of the diffracted beam is maximised. In this case, it can be seen from equation 2.1 that, $\eta_{max} \approx v^2$ provided that $v \ll 1$. Furthermore, for small angles of incidence and for grating vectors nearly perpendicular to the z axis:

$$\eta_{max} = \sin^2 \left(\frac{\pi \Delta n d}{\lambda_r \cos B} \right) \approx \left(\frac{\pi \Delta n d}{\lambda_r \cos B} \right)^2 \quad (2.4)$$

This indicates that the peak diffracted amplitude will increase as the square of both the hologram thickness and the refractive index modulation. For the case of a BSO-based correlator, the relatively small values of refractive index modulation will make this approximation valid.

2.1.1 Refractive Index Modulation of BSO

In order to calculate the peak diffraction efficiency expected from the BSO crystal, the expected values of index modulation need to be deduced. If an electric field of magnitude E exists at some point in the BSO crystal then the resulting refractive index change is given by:

$$\Delta n = \frac{1}{2} n^3 r_{eff} E \quad (2.5)$$

As described in the introduction, such electric fields can be brought about by a separation of positive and negative charge in the BSO. This charge separation will occur if the crystal is illuminated with light possessing a cosinusoidal intensity variation of the form:

$$I = I_0(1 + M \cos(2\pi x/\Lambda)) \quad (2.6)$$

where x is the spatial coordinate in the direction of the grating vector. For this intensity profile [1.73] gives the magnitude of the space charge field as:

$$|E_{SC}| = m \left(\frac{E_A^2 + E_D^2}{(1 + E_D/E_S)^2 + (E_A/E_S)^2} \right)^{\frac{1}{2}} \quad (2.7)$$

where E_A is the magnitude of an external electric field applied to the crystal. E_D is a constant known as the diffusion field, it is independent of material but inversely proportional to the grating spacing.

$$E_D = \frac{2\pi}{\Lambda} \left(\frac{K_B T}{q} \right) \quad (2.8)$$

where K_B is the Boltzman constant, T the absolute temperature and q the charge on the electron. E_S is known as the saturation field and is proportional to the grating spacing. This parameter takes account of the fact that the maximum field will be limited by the number of traps available. The remaining parameter in equation 2.7 is m . This is the reduced modulation ratio and is less than and proportional to

the modulation, M , of the incident intensity pattern in equation 2.6. The reduction arises from the fact that electrons are excited by thermal processes, even in regions of the crystal that are not illuminated. This thermal excitation has the effect of reducing the degree of charge separation.

A convenient method for measuring E_S is derived from the behaviour of equation 2.7 as the fringe spacing is varied. As has already been stated, E_D is inversely proportional to grating spacing while E_S is directly proportional to it. It is therefore convenient to substitute the expressions $E_S = \alpha\Lambda$ and $E_D = \beta/\Lambda$ into equation 2.7. If the resulting equation is differentiated with respect to grating spacing, it is found that, for the case of no applied field ($E_A = 0$):

$$\frac{d|E_{SC}|}{d\Lambda} = \sqrt{m} \left(\frac{\alpha\beta^2 - \alpha^2\beta\Lambda^2}{(\alpha\Lambda^2 + \beta)^2} \right) \quad (2.9)$$

The grating spacing at which this expression is equal to zero is known as the critical fringe spacing, Λ_C , given by $\Lambda_C = \sqrt{\beta/\alpha}$. The second differential at this point is also seen to be negative, indicating a maximum. At this point the modulation of the space charge field is a maximum and $E_S = E_D = \sqrt{\alpha\beta}$, i.e. the saturation and diffusion fields are equal. The value of space charge field at this point is known as the critical field E_C . By setting up a two beam coupling experiment, the fringe spacing that maximises the coupling, Λ_C , can be measured. Since, from equation 2.8, β is a collection of fundamental physical constants, the second constant of proportionality, α , can be calculated from the measured value of Λ_C .

In order to illustrate the effect of the applied field on $|E_{SC}|$, figure 2.2, shows $|E_{SC}|$ plotted against fringe spacing for applied fields of 0, 1, 2 and 4 kVcm⁻¹. To produce the curves, m was taken as 1.0, E_D as $1.57 \times 10^{-6}/\Lambda$ kVcm⁻¹ and Λ_C as $0.5\mu\text{m}$. As expected, the plot for $E_A = 0$ shows a peak at the critical fringe spacing of $0.5\mu\text{m}$. The important point to note is that, for fringe spacings greater than Λ_C , the space charge field increases with applied field. Thus, the diffraction efficiency of any holograms formed in the BSO can be increased by the application of an external field which is a very well known result.

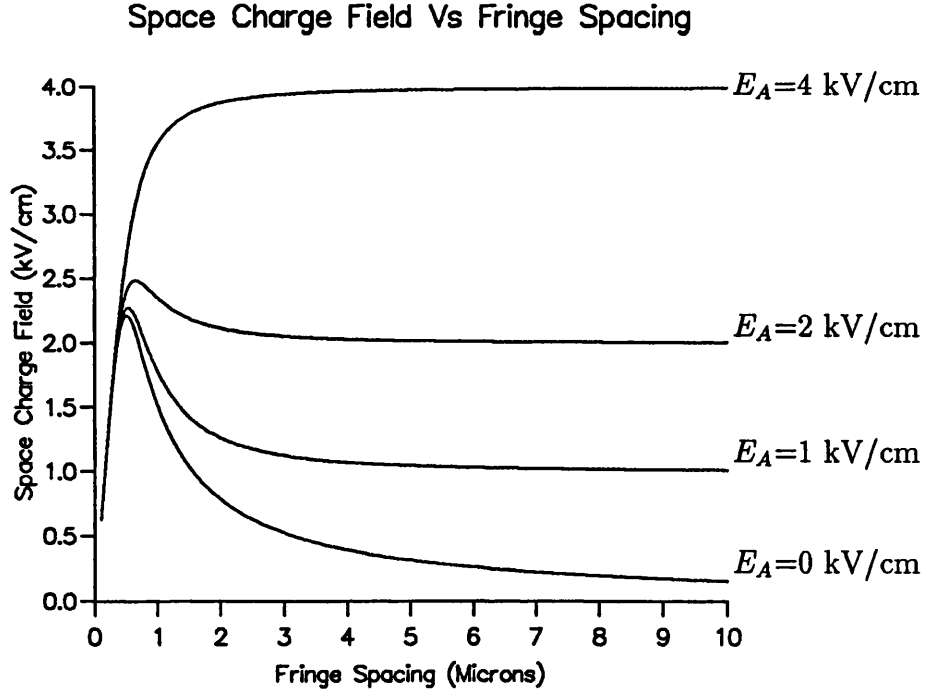


Figure 2.2: $|E_{SC}|$ varying with fringe spacing for various applied fields.

Furthermore, at sufficiently large fringe spacings, the increase in the saturation field means that equation 2.7 reduces to $|E_{SC}| \approx E_A$ and thus the space charge field becomes independent of fringe spacing. This variation in behaviour with fringe spacing will be important in the correlator design since a range of fringe spacings will be produced by the writing beams.

We now have, from equation 2.1, an expression for the peak diffraction efficiency of a hologram of thickness d with a refractive index modulation of Δn . The expected refractive index modulation can be derived from equation 2.5, knowing the magnitude of the electric field present in the crystal. For the case in which the crystal is illuminated with a cosinusoidal grating, the expected electric field magnitude is given by equation 2.7. If the grating spacing is sufficiently large, the field within the crystal will be approximately equal to the externally applied field. The fringe spacing that has to be exceeded for this to be true is the critical fringe spacing, Λ_C . This can be determined by a relatively straight forward experiment to measure the two beam coupling strength, as a function of grating spacing, in the absence of an external field. This same experiment will also allow the determination of the various electric

field terms of equation 2.7. However, two other effects must be taken into account: namely the absorption of light by the BSO; and the Fresnel reflections at each of the two crystal surfaces. As will be discussed in chapter 3, the decision was taken to read out the hologram with light of 633nm wavelength rather than the 514nm that would be used to form the holograms. Values for the absorption coefficient and the refractive index of BSO are:

Wavelength, λ	633nm	514nm
Absorption coefficient, α	1.0cm^{-1}	2.5cm^{-1}
Refractive index, n	2.54	2.61

For the 1mm thick BSO crystal at near-normal incidence these imply total transmissions of 60% at 633nm and 50% at 514nm.

For the BSO sample to be used in the correlator, the electric field was applied in the $\langle 110 \rangle$ direction and the relevant electro-optic coefficient was r_{41} . Experiments were carried out in order to measure the value of r_{41} for the BSO crystal. These gave:

$$r_{41} = (3.6 \pm 2) \times 10^{-12} \text{mV}^{-1} \quad (2.10)$$

This compared well with measurements by Foote [1.74] but is lower than the commonly accepted value of $5 \times 10^{-12} \text{mV}^{-1}$ or the value of $4.25 \times 10^{-12} \text{mV}^{-1}$ by Bayvel [2.3].

In order to measure the critical fringe spacing, a two beam coupling experiment was performed with no externally applied electric field. The coupling was measured as the ratio of the intensity of a weak probe beam with and without the presence of a strong pump (the beam ratio was $\approx 1000 : 1$). The results are shown in figure 2.3 together with a theoretical curve for $|E_{SC}|$. The theoretical curve was generated using a critical fringe spacing of $0.55\mu\text{m}$ and was scaled in order to match the peak values. The overlap of the experimental and theoretical data confirmed that the critical fringe spacing was approximately $0.5\mu\text{m}$. This corresponds to a critical field of 3kVcm^{-1} .

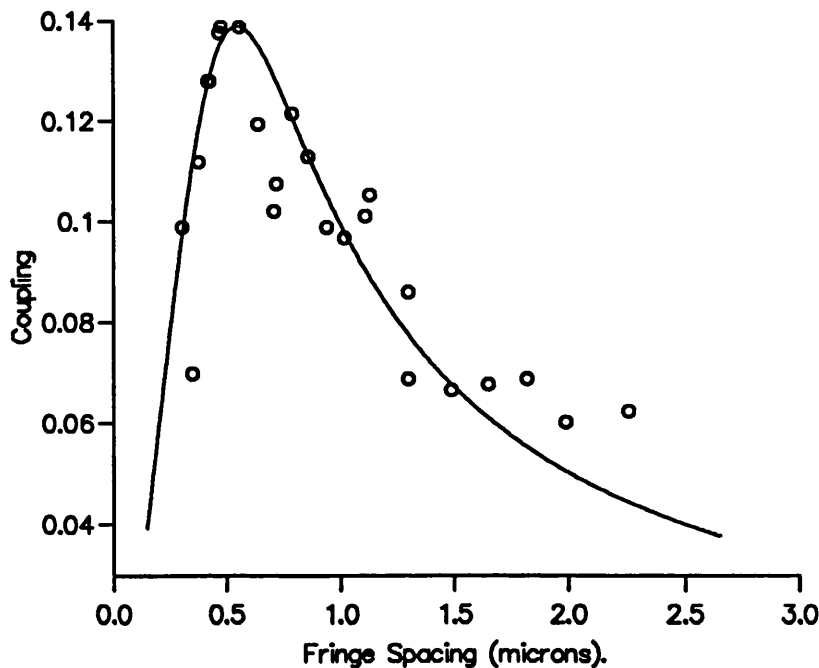


Figure 2.3: Coupling in two wave mixing as a function of fringe spacing. Circles show experimental data while the solid line shows a normalised theoretical curve for $\Lambda_C=0.55\mu\text{m}$.

Thus, assuming that the grating spacings are significantly larger than $0.5\mu\text{m}$, the expected value for the maximum diffraction efficiency will be:

$$\eta = \left(\frac{\pi}{\lambda_r}\right)^2 \left(\frac{1}{2}n^3 r_{41}\right)^2 e^{(-\alpha d)} \left(\frac{4n}{(1+n)^2}\right)^2 d^2 E_A^2 = 1.27 \times 10^{-4} d^2 E_A^2 \quad (2.11)$$

where d is measured in cm and E_A in kVcm^{-1} . This assumes that the refractive index modulation is small as are all the beam angles. The second and third terms are the expressions for absorption and Fresnel transmission at two surfaces respectively.

2.2 Variations in Diffraction Efficiency

The BSO hologram in the correlator architecture will be both read and written by beams which subtend a cone of angles. In order to investigate the effects of varying beam angles, graphs may be plotted from equation 2.1.

2.2.1 Changes in Read-Beam Angle

The first thing to be considered is the effect of changing the read-beam angle away from the optimum Bragg condition. Figures 2.4 and 2.5 show the variation of diffraction efficiency with read-beam angle. The angle is measured relative to the Bragg angle of the hologram, given by equation 2.3. Figure 2.4 shows three plots for holograms of different thicknesses with a constant fringe spacing of $7\mu\text{m}$. Figure 2.5 shows the effects of varying fringe spacing for a 1mm thick hologram. It should also be noted that all beam angles are measured within the holographic medium. A value of $E_A=5\text{kVcm}^{-1}$ was used to determine the maximum diffraction efficiency for the hologram. The graphs illustrate the quadratic increase in peak diffraction with hologram thickness, as well as the increase in angular selectivity for the thicker holograms. At an angle of 0.75° , for example, it can be seen that there is actually greater diffraction from the thinner hologram compared to the thicker ones. Thus, there will be a trade off between peak height and angular selectivity and this is dealt with later. In the case of varying fringe spacing, the three holograms are seen to produce the same peak efficiency, each at its own particular Bragg angle. An increase in fringe spacing leads to a decrease in the angular selectivity of the hologram. An often used parameter is the half-power angle, i.e. the Bragg mismatch that causes the diffraction efficiency to fall to half its maximum value. This will be approximated [1.75] as:

$$2\Delta B_{\frac{1}{2}} \approx \frac{\Lambda}{d} \quad (2.12)$$

Some sample calculations, using the half-power approximation, have shown that it gives a value of $0.4\eta_{max}$, and not $0.5\eta_{max}$, when substituted into equation 2.1. However, this only makes a significant difference for the broader selectivity functions. If the maximum Bragg miss-match is known, then the hologram thickness may be optimised to produce the maximum diffraction efficiency when this miss-match is present. This is important since it is the weakest diffraction signal present in the system that will ultimately limit the performance. Two competing effects are present: firstly, the maximum diffraction efficiency increases with thickness; and secondly, there is a decrease in the minimum to maximum efficiency ratio.

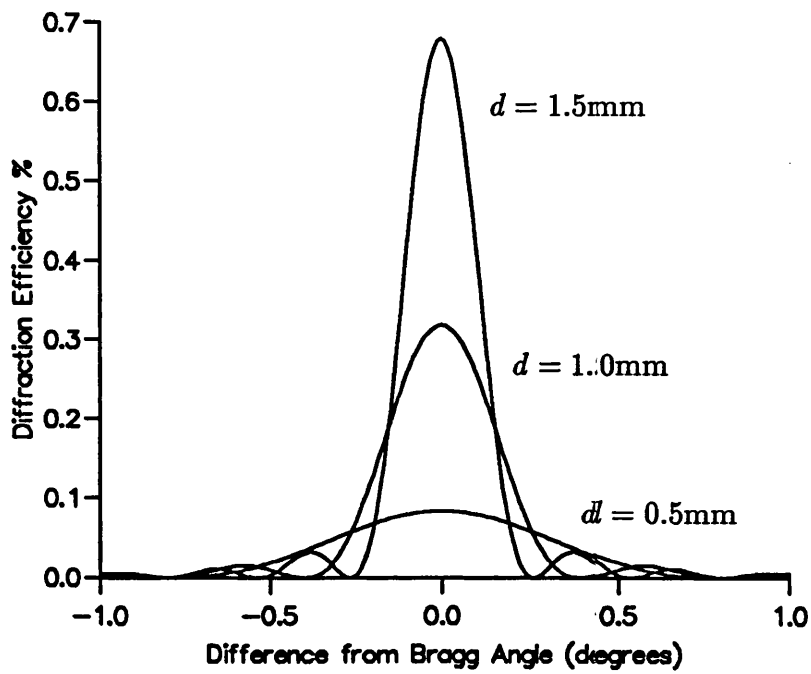


Figure 2.4: Diffraction efficiency vs read-beam angle at a constant fringe spacing but for different thicknesses of grating.

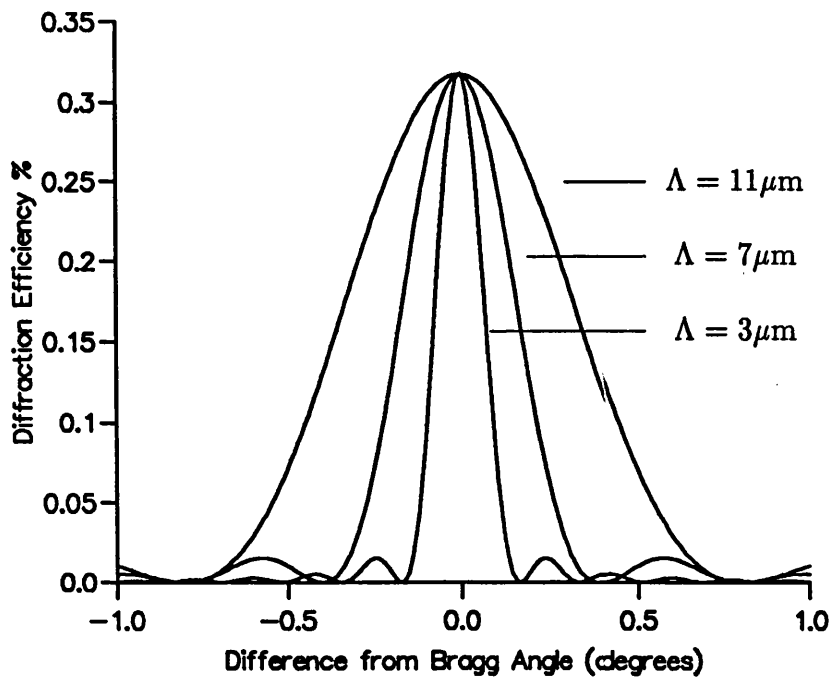


Figure 2.5: Diffraction efficiency vs read-beam angle for different fringe spacings but with a constant thickness of $d = 1\text{mm}$

Equation 2.1 can be rewritten:

$$\eta = \frac{\sin^2 \sqrt{(a^2 + b^2)d^2}}{1 + \frac{a^2}{b^2}} = \frac{\sin^2(cd)}{e} \quad (2.13)$$

Where d is the hologram thickness and:

$$\begin{aligned} a &= \pi \Delta n / (\lambda_r \sqrt{c_{RCs}}) \\ b &= \Delta B \pi \sin(\phi - B) / (\Lambda c_s) \end{aligned} \quad (2.14)$$

To maximise η with respect to the hologram thickness the expression above may be differentiated with respect to d and the result equated to zero.

$$\frac{\partial \eta}{\partial d} = \frac{c \sin(2cd)}{e} = 0 \quad (2.15)$$

Ignoring the solution $d = 0$, the above expression is true when the argument of the sine function is a multiple of π . Hence, the optimum thickness which maximises the diffraction efficiency for a given deviation from the Bragg angle is given by:

$$d = \frac{\pi}{2\sqrt{a^2 + b^2}} \quad (2.16)$$

Thus, the diffracted power at the edge of the field of view of the read-beam can be maximised at the cost of increased variation across the input field. It should be noted that the hologram is only strongly selective in the direction parallel to the grating vector and movements of the read-beam in a direction perpendicular to this have relatively little effect [1.62]. In order to remove the angular effects on the read-beam, it would be possible to place a blocking filter in the system, with a transmission profile which is the inverse of the angular selectivity function. In this case, the angular selectivity would reduce the apparent peak diffraction efficiency, since this value would be reduced to that produced at the edge of the field of view. An even better solution would be to use beam shaping holograms to change the intensity profile of the incident laser beam [2.4] in order to counteract the selectivity effects. In theory, this could greatly improve the light utilisation of the system. A third solution would be to perform a simple digital operation on the captured correlation plane, to boost the intensity values at its edges. This, however, would lead to different signal to noise ratios across the output plane unless the output detector had a sufficiently

large dynamic range. All things considered, the transmission mask would seem to be the simplest effective solution.

2.2.2 Changes in Write-Beam Angle

The second case to consider is that in which the angles of the writing beams are varied. It will be assumed that the read-out beam has been Bragg matched to a hologram formed by two plane writing beams of free-space wavelength λ_w . However, this may not necessarily be equal to the read-out wavelength. This initial hologram is then altered by changing the angles of the writing beams whilst the read-beam angle remains fixed. The two writing beams initially interfere in the medium, with a relative angle of 2α , as shown in figure 2.6. The fringes, being lines of constant phase, are shown as dotted lines.

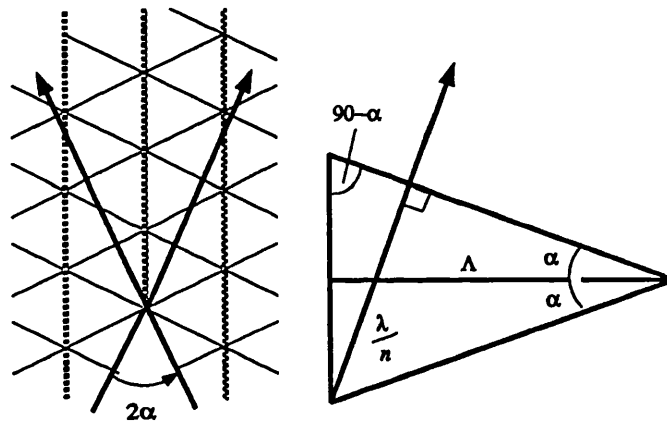


Figure 2.6: Grating formed by two interfering beams.

Figure 2.6 also shows an expanded view, from the geometry of which the fringe spacing, Λ , may be calculated.

$$\Lambda = \frac{\lambda_w}{2n \sin \alpha} \approx \frac{\lambda_w}{2n\alpha} \quad (2.17)$$

Note that, it is again assumed that all beams are in the holographic medium and so the wavelength is λ_w/n .

The grating vector \mathbf{K} is perpendicular to the fringe planes, i.e. perpendicular to the angle bisector of the two writing beams. For this study the initial conditions will

always be with the bisector lying along the z axis, i.e. with $\phi = \pi/2$ in figure 2.1. If the read-out beam now remains stationary whilst the angle of one of the writing beams is changed, this will effect the intensity of the diffracted beam. To quantify the actual effects produced, the system geometry shown in figure 2.7 is considered.

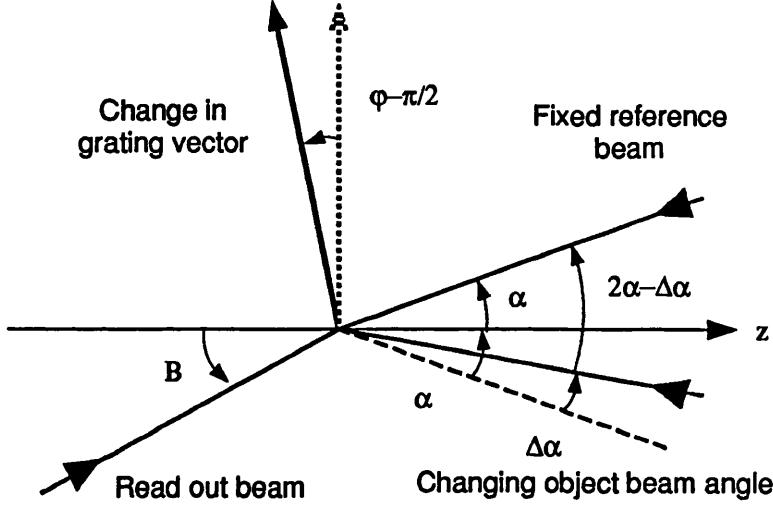


Figure 2.7: Varying the writing object beam angle with fixed reference and read-beams.

The writing beam that is approximately counter propagated by the read-beam has been named the 'reference beam', whilst the other is designated the 'object beam'. This is done in anticipation of the Vander Lugt correlator architecture which the holographic system will be part of. Initially, with $\Delta\alpha=0$, the grating spacing is given by equation 2.17 and the fringes are parallel with the z axis. In order to Bragg match the read-out beam to this grating, it is set at an angle of $B_{(\Delta\alpha=0)}$. Note that, if $\lambda_w=\lambda_r$ then $B = \alpha$, i.e. the read-beam counter propagates the reference beam.

$$B_{(\Delta\alpha=0)} = \sin^{-1} \left(\frac{\lambda_r}{2\Lambda n} \right) \approx \frac{\lambda_r \alpha}{\lambda_w} \quad (2.18)$$

First, consider the effect of changing the angle of the object beam by an amount $\Delta\alpha$; it is conventional to denote anti-clockwise rotations as positive. The angle between the two beams becomes $2\alpha - \Delta\alpha$ since, for a positive value of α , the angle is reduced. Two affects are observed; firstly the grating fringe spacing changes and secondly the fringe orientation changes (a fact that was omitted from the brief analysis

of [1.48] but noted by [1.62]). The grating fringe spacing becomes:

$$\Lambda \rightarrow \frac{\lambda_w}{2n \sin[(2\alpha - \Delta\alpha)/2]} \approx \frac{\lambda_w}{n(2\alpha - \Delta\alpha)} \quad (2.19)$$

It is important to remember that if the fringe spacing becomes too large, the grating will cease to constitute a thick hologram and the analysis will break down. The grating vector now makes a different angle with the z axis, given by:

$$\begin{aligned} \phi_{(\Delta\alpha)} &= \alpha - (2\alpha - \Delta\alpha)/2 + \pi/2 \\ &= \frac{\Delta\alpha}{2} + \frac{\pi}{2} \end{aligned} \quad (2.20)$$

Note that anti-clockwise rotation of the object beam causes an anti-clockwise rotation of the fringes. Now the Bragg condition given by equation 2.3 must be satisfied with respect to the new values of Λ and ϕ given by the two previous equations.

$$\cos\left(\frac{\Delta\alpha}{2} + \frac{\pi}{2} - B_{\Delta\alpha}\right) = \frac{\lambda_r}{2\lambda_w}(2\alpha - \Delta\alpha) \quad (2.21)$$

For small angles, this rearranges to give:

$$B_{\Delta\alpha} = \frac{\lambda_r}{\lambda_w}\alpha + \frac{\Delta\alpha}{2} \left(1 - \frac{\lambda_r}{\lambda_w}\right) \quad (2.22)$$

which gives a Bragg mismatch of:

$$\begin{aligned} \Delta B_{\Delta\alpha} &= B_{\Delta\alpha} - B_{\Delta\alpha=0} \\ &= \frac{\Delta\alpha}{2} \left(1 - \frac{\lambda_r}{\lambda_w}\right) \end{aligned} \quad (2.23)$$

If ΔB is positive then the read-out beam is displaced in an anti-clockwise direction from the required Bragg angle. Note that if $\lambda_w = \lambda_r$ then ΔB is zero for all object beam angles, which is a well known result. In this case the effects of changes in the fringe angle and the fringe spacing exactly cancel. It is interesting to note that for write and read wavelengths of 514 and 633nm respectively, the Bragg mismatch produced is only about one tenth of the actual angular change in the object beam angle. If the expression of equation 2.12 is considered, this in turn suggests that the angle associated with a drop in power to the half-peak value will be ten times that associated with the read-beam. Figure 2.8, which was computed using a full

unapproximated form of equation 2.23, illustrates the change in diffraction efficiency with object beam angle. It assumes writing and reading wavelengths of 514nm and 633nm respectively and shows the effects of varying BSO thickness. The read-beam is fixed at the angle required to Bragg match it to an initial $7\mu\text{m}$ grating and the object beam is then moved towards (positive $\Delta\alpha$) and away from (negative $\Delta\alpha$) the reference beam. As expected the actual sensitivity of the system to changes in object beam angle is less than that for read-beam angles. In the case of figure 2.8 it would be expected that a change in read-beam angle of only 0.2° would have caused the diffracted signal to fall to half its peak value.

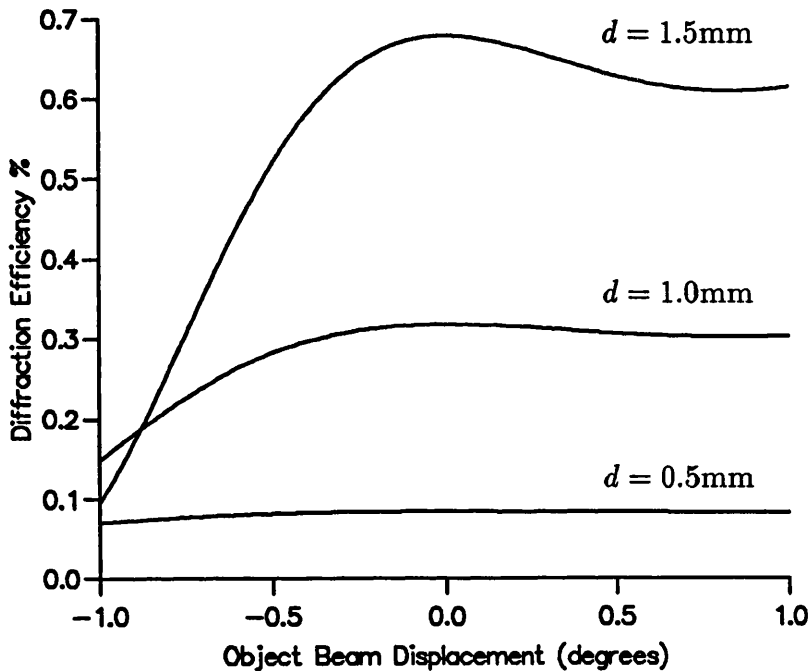


Figure 2.8: Diffraction efficiency varying with object beam angle for a system initially set up with $\Lambda = 7\mu\text{m}$

The plot of figure 2.8 is asymmetric; reducing the angle between the two writing beams has less effect than the corresponding increase. This is due to the higher angular tolerance of large fringe spacing holograms, with Λ appearing in the denominator of ξ in equation 2.1. If higher initial fringe spacings are used, the symmetry of the curves improves. This is due to the fact that, at higher fringe spacings, the relatively small angular changes needed to reduce the diffraction efficiency to zero

will only cause a small percentage change in grating spacing.

If, instead of the object beam, the reference beam angle is changed then the fringe spacing will become:

$$\Lambda \approx \frac{\lambda_w}{2\alpha + \Delta\alpha} \quad (2.24)$$

This is the same expression as that derived previously but now a positive value of $\Delta\alpha$ leads to a decrease in fringe spacing. The angle required for Bragg matching to the altered grating spacing is the same as before and so the Bragg mismatch ΔB is now given by:

$$\Delta B_{\Delta\alpha} = \frac{\Delta\alpha}{2} \left(1 + \frac{\lambda_r}{\lambda_w} \right) \quad (2.25)$$

This expression is identical to equation 2.23 except for a change in sign of the bracketed term. For a given value of $\Delta\alpha$:

$$\frac{\Delta B_{\Delta\alpha}(\text{Ref.})}{\Delta B_{\Delta\alpha}(\text{Obj})} = \frac{\lambda_r + \lambda_w}{\lambda_r - \lambda_w} \quad (2.26)$$

For the case when $\lambda_r = 633\text{nm}$ and $\lambda_w = 514\text{nm}$, a rotation of the reference beam produces a Bragg mismatch that is about ten times worse than that caused by an equal rotation of the object beam. This means that the effect of a given angular change in the reference beam would be roughly equal to the equivalent change in the read-beam. This is important if one considers a joint transform correlator in which two images, each with a large field of view, are being cross-correlated. In this case, one of the images will suffer from the same angular selectivity problems as an image imposed on the read-beam in the Vander Lugt architecture. Thus, for large images, there is nothing to choose between the two architectures from the point of view of angular selectivity.

If the hologram is being interrogated by a read-out beam of the same wavelength as the writing beams, then changes in the object beam angle have no effect. Figure 2.9 illustrates the case for a grating with an initial $2\mu\text{m}$ fringe spacing (which intensifies the angular selectivity effects). The hologram is 1mm thick and three curves are plotted for read-beam wavelengths of 514, 570 and 633nm. Thus, if it is required to correlate a large field of view image with a small reference, the ideal solution would be a system with 514nm read and write beams. The large field of view image could

then be placed on the writing beam with minimal angular selectivity effects being observed. However, if the image on the read-side is large, then it will produce high intensities in the Fourier plane which will, in turn, perturb the hologram.

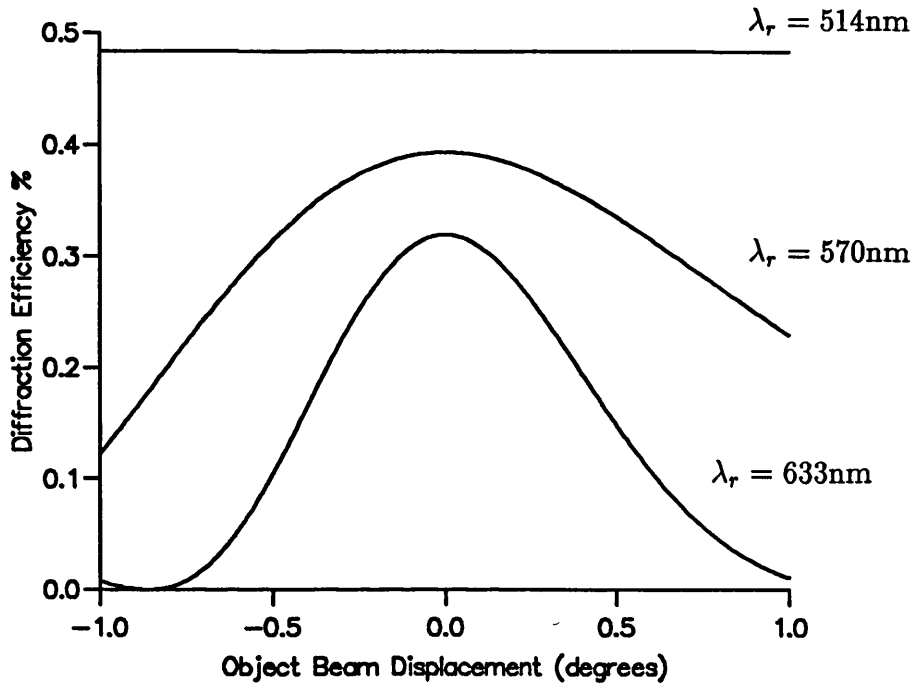


Figure 2.9: Diffraction vs object beam angle for a 1mm thick hologram with an initial fringe spacing of $2\mu\text{m}$

When considering the case of changes in the read-beam angle, it was shown that the effects could be negated by the use of a spatially varying intensity filter. In the case of the writing beams however, this solution is not appropriate. In this case, the strength of the signal due to a particular part of the image can only be manipulated by altering the modulation of the hologram fringes formed by it. However, every point in the Fourier transform plane potentially contains information from every point in the image; so no filtering can be applied in this plane. If any attempt is made to alter the intensities in the image plane on the write side of the correlator, the results will be unpredictable. This is because the act of reducing the intensity of one part of the image will merely alter the part of the Fourier spectrum that the reference beam is balanced to: a fact that was overlooked by [1.62]. Altering the intensities in the image plane, may only have the desired effect in the case when the

reference beam is balanced to the d.c of the Fourier transform. This will not usually be the case since the unweighted correlation functions thus produced are of little use for image processing.

It would seem therefore that there is little that can be done as far as correcting the problems due to an angular spread in the writing object beam. However, the relative reduction in the angular selectivity on the writing side of the hologram suggests that this may not pose a serious problem for the proposed system. If the problem was found to be significant, then an alternative approach to producing edge enhanced correlation could be used. One such approach could use a reference beam intensity that was balanced to the d.c. of the Fourier transform. In this case, edge enhancement could be performed by using an incoherent erase beam to remove the area around the d.c.

2.2.3 Beam Angles Produced in the Correlator

Given expressions for the variation of diffraction efficiency with beam angle, it remains to derive expressions for the angles that will be found in the correlator. In order to do this, a ray matrix formulation is used. Four system parameters are considered: the focal length of the Fourier transform lens; the physical size of the input image; the spatial frequencies associated with it; and its distance from the lens.

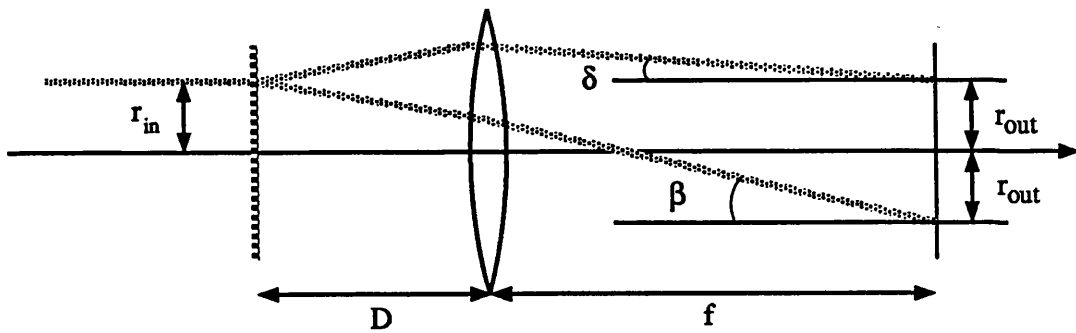


Figure 2.10: Fourier transform of a grating image.

Figure 2.10 illustrates a lens which is forming the Fourier transform of a grating

which has an amplitude transmission of:

$$T_a = \sin(2\pi x/u) \quad (2.27)$$

where u is the spatial frequency of the grating. The grating is an example of a simple image comprising only a single spatial frequency. The images fed into the correlator will obviously contain many frequencies but the CRT/SLM combination will limit the maximum frequency present. The grating is probed by a relatively small laser beam parallel with the optical axis of the system. This small probe beam allows an examination of the change in beam angles in the Fourier plane as a function of position in the input plane. The 'image' is at some arbitrary distance D from the lens; and the plane of interest is the lens' back focal plane. On passing through the grating, the beam is split into two diffracted orders, at angles of $\pm\lambda u$ to the optical axis. In order to calculate the angle and the distance from the axis of this ray in the Fourier plane, standard ray matrix notation (chapter 6 of [3.2] for example) may be used.

$$\begin{pmatrix} \theta_{out} \\ r_{out} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ f & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ D & 1 \end{pmatrix} \begin{pmatrix} \theta_{in} \\ r_{in} \end{pmatrix} \quad (2.28)$$

On performing this matrix multiplication, and making the substitution $\theta_{in} = \pm\lambda u$, it is found that:

$$\begin{aligned} r_{out} &= f\lambda u \\ \delta &= -\frac{r_{in}}{f} - \lambda u \left(1 - \frac{D}{f}\right) \\ \beta &= -\frac{r_{in}}{f} + \lambda u \left(1 - \frac{D}{f}\right) \end{aligned} \quad (2.29)$$

The expression for r_{out} is the well known result for the scaling of the Fourier transform. It is important to note that if $D = f$ then the angles of the rays in the Fourier plane are determined *solely by the position from which they originated*. In this case the magnitude of the angular spread depends only on the size of the input image and the focal length of the lens. In the correlator architecture, the optical axis of the system shown in figure 2.10 will be Bragg matched to the BSO hologram. Any ray with an output plane angle of zero, as given by equation 2.29, will be diffracted with maximum efficiency. Any change in angle will lead to a drop in diffraction for

the particular part of the image and/or spatial frequency considered. It is important to note that the angles produced by the lens will be multiplied by a factor of $1/n$ on entering the holographic medium.

2.3 Validity of the Use of Coupled Wave Theory

The coupled wave theory is very useful for predicting, in a conveniently expressed form, the behaviour of a thick holographic element in the correlator. It is important to assess those conditions under which the theory may become invalid; and what the consequences of such failure may be as regards the design process.

In the derivation of the coupled wave theory it was assumed that the only waves propagating in the hologram were the read-beam and one diffracted beam. In this case the hologram could be described as 'thick'. However, as the fringe spacing increases, or as the hologram thickness is reduced, more diffracted orders will be seen. Eventually the alteration of the hologram parameters will produce an output which is seen to be the Fourier transform of the refractive index modulation-profile. In this case the hologram can be described as thin. The thin hologram possesses the important property that its diffraction efficiency is independent of read-beam angle. In [1.75], Kogelnik put forward a grating parameter Q , which was originally considered in the field of acoustic diffraction.

$$Q = \frac{2\pi\lambda_r d}{n\Lambda^2} \quad (2.30)$$

It was suggested in [1.75] that the coupled wave theory will give good results when $Q \gg 1$, and that it begins to give good results when $Q = 10$. However, Kaspar [2.2], in his extension of Bruckhardt's work, suggested that the coupled wave theory will give good results for low Q values *when the refractive index modulation is small*. Since, for the BSO crystal, it is expected that this will be the case, this finding tends to support the use of coupled wave theory for the correlator design even if Q falls as low as 1.3 [2.2].

Further support for the assumption that low refractive index modulation extends the usefulness of the coupled wave theory can be found by considering the diffraction

efficiency predicted for thin holograms. In particular it is the thin phase grating that is of interest. In order to predict the strength of the first-order diffracted beam from a thin hologram, the analysis presented by Goodman on page 69 of [1.5] is used. In this reference the transmittance function of a phase grating is Fourier transformed and expressions for the maximum intensities in the far field pattern are obtained. The transmittance function used in [1.5] is:

$$t(x, y) = \exp \left[i \frac{m}{2} \sin(2\pi f_o x) \right] \text{rect} \left(\frac{x}{l} \right) \text{rect} \left(\frac{y}{l} \right) \quad (2.31)$$

In this expression, m represents the peak-to-peak excursion of the phase delay and f_o is the frequency of the grating. To match this to the parameters used in the coupled wave theory, the identity given below may be used.

$$m = \frac{4\pi \Delta n d}{\lambda} \quad (2.32)$$

This is the phase delay introduced as the light travels through a thickness d , with maximum and minimum refractive indices of $n \pm \Delta n$. Another important difference between the notation in the two theories is the inclusion of boundary conditions in the thin hologram calculations. Unlike the coupled wave theory, the thin hologram is contained in a square aperture of side l . In order to take this into account, a square aperture may be imposed on the diffracted order of the thick hologram. This is done by considering the diffraction pattern formed by a square aperture, illuminated by a plane wave with an intensity predicted from the diffraction efficiency for the thick hologram.

The peak intensity in the q th order produced at a distance z from the thin phase grating is given in [1.5] as:

$$\eta[\text{thin}]_{\max} = \left[\frac{l^2}{\lambda z} J_q(m/2) \right]^2 \quad (2.33)$$

Where J_q is a Bessel function of the first kind, order q . Note that in order to be valid this expression assumes $z \gg l$. To compare this with the coupled wave theory result, the far field pattern of a square aperture of side l illuminated by a plane wave of intensity $\eta[\text{thick}]_{\max}$ (from equation 2.4), must be considered. The peak intensity

of the far-field pattern can be simply obtained from the expression for the intensity pattern given in [1.5] as:

$$I_{max} = I_{in} \left(\frac{l^2}{\lambda z} \right)^2 \quad (2.34)$$

where I_{in} is the intensity incident on the aperture. By replacing I_{in} with the expression for $\eta[thick]_{max}$ from equation 2.4, and substituting for m into the equation for $\eta[thin]_{max}$, the following expression is obtained for the ratio of intensities for diffracted orders from thin and thick holograms:

$$\frac{\eta[thin]_{max}}{\eta[thick]_{max}} = \left(\frac{J_1(2\pi\Delta nd/\lambda)}{\sin(\pi\Delta nd/\lambda)} \right)^2 = \left(\frac{J_1(2x)}{\sin(x)} \right)^2 \quad (2.35)$$

But, for small x , the following approximations are valid:

$$J_1(2x) \approx x \text{ and } \sin(x) \approx x \quad (2.36)$$

x will be small if the refractive index modulation, Δn , of the grating is small and in this case:

$$\eta[thin]_{max} \approx \eta[thick]_{max} \quad (2.37)$$

Thus, the two theories, derived in totally different ways, give the same peak diffracted signal. In fact, with a hologram thickness of 1mm., Δn values of up to 7×10^{-5} lead only to a 10% difference in the diffraction efficiencies predicted by the two theories. This equality implies that, in figure 2.8, the diffraction efficiency should climb back towards the peak Bragg matched value at some point before the object and reference beams become colinear. Since the hologram will then be thin, the diffraction efficiency will subsequently be independent of beam angle. As there is no predicted drop in diffraction efficiency, it will be beneficial to use holograms as close to the thin regime as possible.

In the derivation of coupled wave theory it was assumed that the fringe profile was cosinusoidal. However, the predictions of the BBTM suggest that this may not be the case if the modulation of incident intensity fringes approaches 1. In this case the profile of the refractive index grating may be far from the idealised form used in the prediction of the diffraction efficiencies of the thick and thin holograms.

Figure 2.11 is a sketch of the fringe profiles taken from [1.73] and it shows the case for a photorefractive material with an applied field.

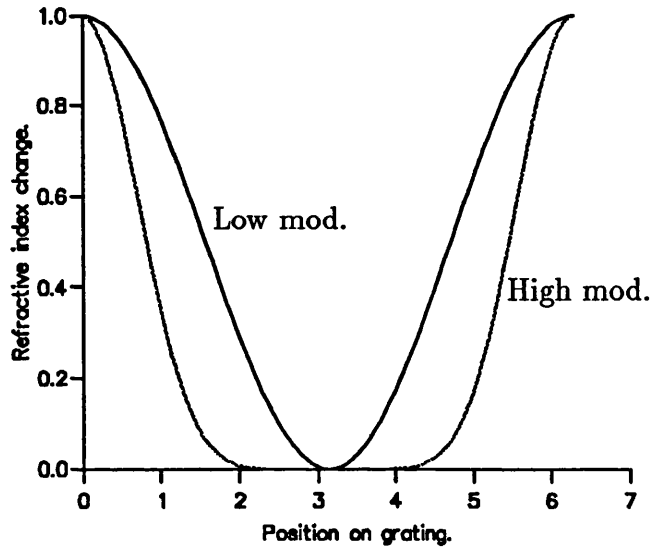


Figure 2.11: Fringe profile for different grating intensity modulations.

It is outside the scope of this work to predict the precise effect of this alteration in fringe profile on the coupled wave theory. It is usually assumed that only the fundamental Fourier component of the grating interacts with the incident light. This assumption is based on Bragg matching considerations. The modified fringes of figure 2.11 can be considered as a Fourier sum of a large number of different cosinusoidal gratings. However, it is assumed that the read-beam is only Bragg matched to one of these gratings, the fundamental. There will be no diffraction from the other components since these, having different periods will require different Bragg matching conditions. If the amplitudes of the other Fourier components became relatively large, then significant signals might be obtained when nearly Bragg matching to these other components. This might appear to broaden the angular selectivity when the amplitudes of other Fourier components, near the fundamental, become large. If the hologram becomes thin, the effects, on diffraction, of the changes in the fringe profile could be predicted by computing its one dimensional Fourier transform. In some experiments, in which a He-Ne laser was diffracted from a high modulation grating, multiple diffraction orders were observed suggesting the non-sinusoidal profile of the fringes.

When predicting how a shift in the object beam angle would affect the diffracted signal, it was assumed that refractive index modulation was independent of fringe spacing. This implies that the electric fields, present in the material, are dominated by the externally applied field, as discussed earlier in this chapter. This will be true only at fringe spacings significantly larger than the critical fringe spacing, Λ_C . As described in the previous section, this fringe spacing was measured as approximately $0.5\mu\text{m}$. From the curves of figure 2.2, it would be expected that, in this case, the diffraction efficiency will remain constant with grating spacing for values higher than about $5\mu\text{m}$. In order to test both this, and the predictions concerning the peak diffraction efficiency, a non-degenerate four wave mixing experiment was carried out.

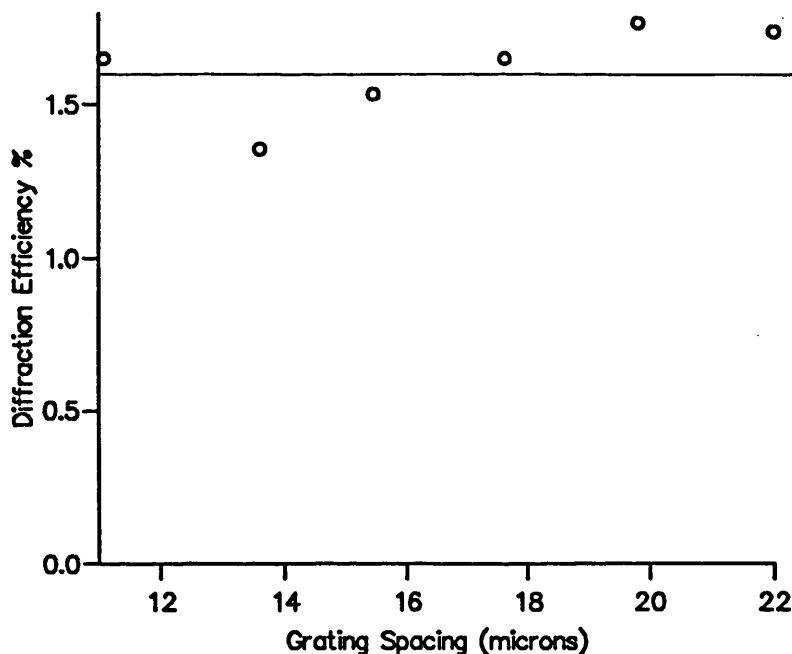


Figure 2.12: Diffraction efficiency varying with fringe spacing for a 1mm thick BSO hologram with an applied field of 5kVcm^{-1} . Diffraction efficiency should be constant (solid line).

Diffraction efficiencies were measured at fringe spacings between 10 and $22\mu\text{m}$. The results were produced by probing a grating, written with an Ar^+ laser, with a beam from a He-Ne laser. The intensity modulation of the writing grating was approximately unity and a field of 7kV was applied across a 14mm wide, 1mm thick BSO crystal. The results are shown in figure 2.12 and have been re-scaled to include

the losses due to absorption and reflections that attenuate the probe beam as it passes through the BSO.

For this BSO sample, a fringe spacing of $12\mu\text{m}$ corresponded to a holographic Q parameter of 10. Thus, for fringe spacings greater than this the hologram may be moving into the thin regime. However, as predicted, the diffraction efficiency remains approximately constant. The average value is 1.7%. However, from equation 2.11 a value of 0.3% would have been predicted. This discrepancy is more than could be accounted for by having underestimated the value of r_{41} . It was thought that it was due to an inaccurate measurement of the actual voltage across the region of the crystal that was supporting the hologram.

The coupled wave theory is derived assuming that the polarisation of the read-beam is perpendicular to the plane of incidence shown in figure 2.1. For the system being considered here, it may be desirable, from signal to noise considerations, to use different input polarisations. Even if, at the front face of the crystal, the polarisation corresponded to that used in the derivations, it would be altered by optical activity and field induced birefringence as the beam passed through the crystal. In an appendix to [1.75] Kogelnik considered the effect of polarisation on the maximum diffraction efficiency for light polarised parallel to the plane of incidence. It was shown that, for a grating vector angle of $\pi/2$, the coupling constant was reduced by a factor of $\cos(2B)$. It is interesting to note that for Bragg angles of up to 9° the reduced coupling constant would still be 95% of the one for perpendicular polarisation. However, it could be possible that polarisation effects decrease the apparent thickness of the hologram as suggested in [1.62]. This will happen if polarisation changes decrease the diffraction efficiency towards the back face of the crystal. Thus, most of the diffracted signal would be produced in a layer of reduced thickness at the front of the hologram. In this case, the reduction in angular selectivity would be accompanied by a reduction in the peak diffracted signal. However, it should be noted that, for a 1mm thick sample of BSO, the rotation of a linearly polarised beam at 633nm will be only about 24° which can be translated to $\pm 12^\circ$ about the optimum angle for diffraction. Also, the maximum phase difference between the o and e waves, due to

the field induced birefringence, will only be 0.16λ .

In figure 2.1 the grating is embedded in a medium of the same average refractive index as the grating itself. In reality, the grating in the BSO will be bordered by two air interfaces. Thus, the grating is inside a Fabry-Perot oscillator. The effects of this were discussed in another paper by Kogelnik [2.5]. The discussion centres on the calculation of the peak Bragg matched efficiency of the grating. It does not specifically include angular matching effects, other than the change in the apparent Bragg angle of the system. In the case of small refractive index modulation, it was shown that the peak efficiency calculated by the coupled wave theory should be modified by:

$$\eta_{max} \rightarrow \eta_{max} \frac{(1 - R)^2(1 + 2R \cos(2\beta d) + R^2)}{(1 - R^2)^2 + 4R^2(1 + \cos^2(2\beta d)) - 4R(1 + R^2) \cos(2\beta d)} \quad (2.38)$$

Where R is the *power* reflectance of the grating boundaries and β , a propagation constant for light in the medium, is given by $\beta = (2\pi n/\lambda) \cos B$. Note that B is still the Bragg angle in the grating medium. By plotting the modifying term against $\cos(2\beta d)$ it can be found that the limiting cases are when the cosine is equal to plus or minus one. For the BSO crystal, R can be calculated from the Fresnel equations and, if n is taken as 2.5, $R = 18\%$ at near-normal incidence. The two limiting cases then give $2.10\eta_{max}$ and $0.23\eta_{max}$ respectively, which represents a significant change in the diffraction efficiency. Note that the effect is very oscillatory in terms of both the Bragg angle, B , and the hologram thickness, d . At near-normal incidence, a change in d of $\lambda/2n$ will cause $\cos 2\beta d$ to change from -1 to 1. This would make the diffraction efficiency very difficult to predict. It was found however, that the BSO crystal was wedge shaped, with a difference in thickness of 50λ between opposite edges. Because of this, the result of the multiple reflections will be fringes across the output, the fringe modulation being about 0.4.

In order to test the angular selectivity predictions of the coupled wave theory, a simple experiment was performed on a 1mm thick sample of BSO. The diffraction efficiency of a He-Ne laser beam was monitored as a function of its angle of incidence, measured with respect to the holographic fringe planes. The efficiency was monitored by a photo-detector placed in the diffracted beam; and care was taken to ensure that

the results were not influenced by variations in the probe beam intensity. A fringe spacing of $10\mu\text{m}$ was used with an applied field of 7kVcm^{-1} . The experimental results are shown in figure 2.13, together with a theoretical curve generated from Kogelnik's equations.

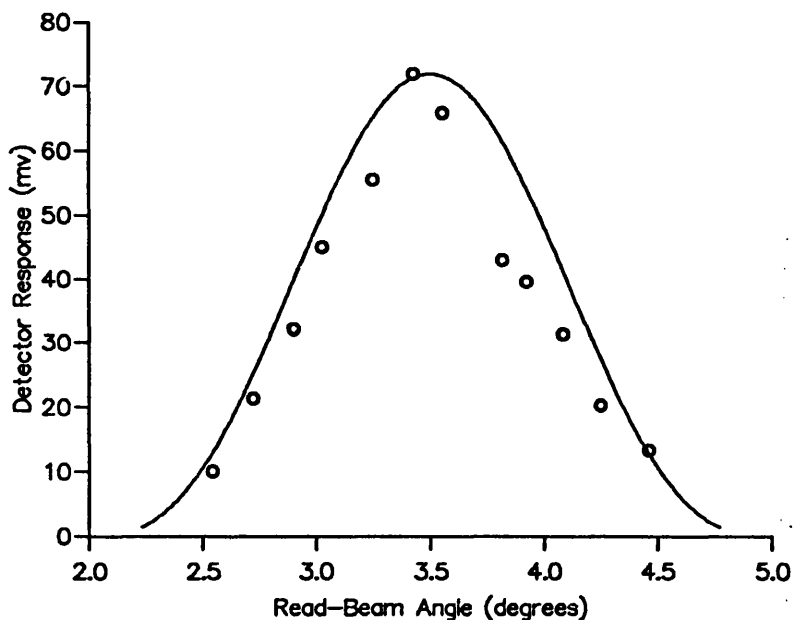


Figure 2.13: Experimental data (circles) and theoretical prediction (curve) for variation of diffraction efficiency with read-beam angle.

The theoretical curve has been rescaled to take account of the fact that the experimental angles were measured outside the BSO. This meant that the actual angular change of the light addressing the holographic grating was about 2.54 times less than that measured outside the crystal. It can be seen from the figure that theory and experiment are in good agreement unlike the results reported by [1.62]. However, it is difficult to assess where the discrepancies between the two sets of results may have arisen.

2.4 Depth of Field in the Fourier Plane

The advantages of using long focal length lenses in the presence of thick holographic media extend beyond the reduction of angular selectivity. The longer the focal length

of the transform lens, the greater will be the 'depth of field' through the Fourier plane. It must be assumed that, for a moderately thick BSO crystal, the hologram will extend through its entire thickness. This is because there will be no significant attenuation of the writing beams as they pass through the crystal. However, the precise Fourier transform is located at a discrete plane and is bordered by Fresnel diffraction patterns on either side of it. An analysis of this depth of field may be found in [3.9]. Here, the through focus intensity pattern for a converging spherical wave, diffracted by a circular aperture, is derived. An expression for the separation of the two zero intensity points either side of the geometric focus, for an aberration free system, is derived. Measurements made on an actual diffraction pattern showed that this separation is misprinted in [3.9], as a factor of 4 is 'lost'. However, from an examination of the text it is found to be:

$$\Delta z = 4\lambda n \left(\frac{f}{a} \right)^2 \quad (2.39)$$

where a is the radius of the aperture. The refractive index, n , was included as it is the depth of field inside the BSO crystal which is important.

2.5 Conclusions

The results of the coupled wave theory are useful for the design of the correlator architecture, because they lead to an analytical solution. Thus, it is relatively easy to make predictions about the behaviour of a thick hologram in the Fourier plane of the system. In general it can be seen that, because of the hologram's angular selectivity, it is important to keep the read-beam as close to the Bragg angle as possible. Thinner holograms, and holograms with larger fringe spacings, will show less angular selectivity. However, the peak Bragg-matched diffraction efficiency will decrease with the square of the hologram thickness. In terms of the read-beam, the angular selectivity effects may be neutralised by simply placing a filter, with an appropriate intensity transmission profile, in the read-beam. This will reduce the diffracted power, across the whole of the output plane, to the value of the minimum diffraction at the edges of the input image. If the maximum angular spread is known, then the

crystal thickness can be optimised to produce the greatest possible diffraction at the limits of the angular spread.

Changing the angle of the write-side object beam will change both the grating spacing and the angle of the grating vector. If the write and read-beams are of the same wavelength these effects exactly cancel. In the correlator architecture, it is expected that wavelengths of 514nm and 633nm will be used respectively to write and read the holograms. Even with this wavelength difference, a change of object beam angle produces only one tenth of the Bragg miss-match that would result from an equivalent change in the read-beam angle. The effect of decreasing the read wavelength is an increase in predicted diffraction and a reduction of angular selectivity. However, it would be accompanied by an increase in the absorption of the read-beam in the BSO which would perturb the hologram. Unlike the case for the read-beam, the angular selectivity, due to changes in object beam angle, cannot be rectified by a simple transmission filter. This is because changes in intensity in the image plane will change the intensity of the corresponding light in the Fourier plane. If weighted correlation functions are being produced, these changes in the Fourier intensities will alter the spatial frequencies to which the reference beam is balanced. Thus, the effect will be to produce different degrees of edge enhancement for different parts of the write-image, and not different apparent intensities. The angular spread of the writing object-beam is not expected to be a problem however, because of the reduced angular selectivity on the write-side.

It is also important to consider the validity of the theory in the context of the design task. The results for a $10\mu\text{m}$ fringe spacing match the theory very closely in terms of variations in efficiency with beam angle. Ultimately, decreasing the hologram thickness, or increasing the fringe spacing, will produce 'thin' holograms for which the theory is not correct. These thin holograms are predicted to have the same efficiency as the thick holograms but without the angular selectivity. Another effect, occurring at small fringe spacings, is a variation of diffraction efficiency with fringe spacing. However, it is not expected that these conditions will be encountered in the correlator.

Other effects that may alter the performance of the hologram are: the effects of non-cosinusoidal fringe profile; polarisation effects; and multiple reflections from the faces of the BSO. A measurement of peak diffraction efficiency produced a result that was five times higher than that predicted. It is thought however, that the most important factor will be the angular effects and not the maximum diffraction efficiency. It is hoped that the coupled wave theory will at least predict the worst possible effects to be encountered. If the actual effects observed in the final correlator system are less severe than expected, this will be of benefit, but it will not be relied upon.

References: Chapter 2

- 2.1 Bruckhardt, C.B., "Diffraction of a Plane Wave at a Sinusoidally Stratified Dielectric Grating", J.O.S.A., **56** p.1502 (1966).
- 2.2 Kaspar, F.G., "Diffraction by Thick, Periodically Stratified Gratings with Complex Dielectric Constant", J.O.S.A., **63** p.37 (1973).
- 2.3 Bayvel, P., "Investigation of the Electro-optic properties of $\text{Bi}_{12}\text{SiO}_{20}$ and $\text{Bi}_{12}\text{GeO}_{20}$ —Application to Sensors", Jour.of.Optical Sensors, **2** p.413 (1987).
- 2.4 Roberts, N.C., "Beam Shaping by Holographic Filters", Appl.Opt., **28** p.31 (1989).
- 2.5 Kogelnik, H., "Bragg Diffraction in Hologram Gratings with Multiple Internal Reflections", J.O.S.A., **57**, p.431 (1967).

Chapter 3

Other Factors Pertinent to the Correlator

Contents of Chapter 3

3.1	Effect of Input Plane Positions	84
3.1.1	The Location of the Correlation Plane	84
3.1.2	Unwanted Spatial Filtering Effects	90
3.2	Lens and Wavelength Requirements	96
3.2.1	Two-Element Compound Lenses	97
3.3	Powers Required in the Read and Write-Beams	108
3.3.1	Correlation Plane Intensities	109
3.3.2	Fourier Plane Intensities	113
3.3.3	Response Time of the BSO	115
3.4	Tilted Fringes in an Applied Field	119
3.5	Limits on Fringe Spacing	122
3.6	Comparison of BSO and Holographic Film	124
3.7	Beam Polarisation	135
3.8	Conclusions	136

Abbreviations/symbols used in Chapter 3

BSO	Bismuth Silicon Oxide (Chapt. 1).
CCD	Charge Coupled Device (Chapt. 1).
CRT	Cathode Ray Tube (Chapt. 1).
d.c.	Zero-spatial-frequency component of FT (Chapt 1.).
FT	Fourier Transform (Chapt 1.).
MSF	Matched Spatial Filter (Chapt. 2).
SLM	Spatial Light Modulator (Chapt. 1).
Ar^+	Argon ion (laser) (Chapt. 1).
D	Separation of two lenses in a compound lens.
E'_A	is the applied field (Chapt. 2) normalised by the critical field.
E_{SC}	Total electric field across BSO.
f	Lens' focal length (Chapt. 2).
I_{FT}	Intensity in the Fourier plane.
k	$2\pi/\lambda$
lpmm^{-1}	Line-pairs per mm (spatial frequency).
$N \times N$	Number of pixels in an image (Chapt. 1).
P_{las}	Laser power.
r_{BSO}	Half-width of clear aperture of BSO crystal.
r_{im}	Half-width of input image aperture.
$T_{1/2}$	Time for grating to decay to give half the maximum diffraction.
u	Spatial frequency.

α	Half-angle between holographic writing beams (Chapt. 2).
Λ	Hologram's fringe spacing (Chapt. 2).
Λ_C	Critical fringe spacing for BSO (Chapt. 2).
λ	Free-space wavelength (Chapt. 2).
λ_r, λ_w	Subscripts to denote reading and writing of hologram.
η	Diffraction efficiency (Chapt. 2).
τ_g	Exponential rate coefficient for growth of field in BSO.
ν	Holographic grating frequency normalised by the critical grating frequency.
τ_c	Dielectric lifetime for BSO.
Θ^2	is a diffusion length parameter.
θ	Tilt of holographic grating vector with respect to the direction of the applied field.

During the course of the design work, several important topics were identified that were not covered by previously published papers. The work that was carried out, in order to fill the various gaps, is reported in this chapter. The effect of the input plane position is considered both from the point of view of the correlation plane position and from that of the beam angles produced in the Fourier plane. Following this, the selection of the read and write-wavelengths and the requirements for the Fourier transform lenses are discussed. It is shown how a lens, of any required focal length, can be constructed from a pair of standard 'off-the-shelf' components. The use of ray-tracing lens analysis for selecting appropriate lens combinations is also discussed.

The powers required in the read and write-beams are dictated by: the sensitivity of the output plane detector; the applied field; the thickness of the BSO crystal; and the required BSO response time. It is shown in this chapter how computer simulation can be used to compute the intensities in both the Fourier plane and the correlation plane.

An effect, not hitherto considered, is that of tilted hologram fringes. If the holographic grating vector becomes miss-aligned, with respect to the applied electric field, this will cause the diffraction efficiency to fall. There will always be some points in the write-side image that will produce these tilted fringes. Expressions for the degree of fringe tilt produced in the correlator are derived and experimental results are presented which measure the severity of the effects.

In the previous chapter, it was shown that the use of large holographic fringe spacings reduces the angular selectivity of the holograms. A limit on the maximum fringe spacing is imposed by the need to separate the correlation plane from the undiffracted beam. At the same time, the minimum fringe spacing must be significantly greater than the critical fringe spacing Λ_C . In this chapter, expressions for the maximum and minimum allowable fringe spacing are derived.

Unlike holographic film, the holographic recording in BSO is much less subject to intensity saturation effects. This alters the way in which the recorded amplitude of a Fourier transform relates to its original amplitude. In this chapter, a

one-dimensional, digital simulation of hologram storage is carried out. This allows conclusions to be drawn about the differences between BSO and holographic plate storage.

One other area included in the correlator design, is the optimisation of the read and write-beam polarisations. Although not examined in detail in this research, the subject is reviewed in this chapter for completeness.

3.1 Effect of Input Plane Positions

In order to construct a compact optical correlator, it is often desirable to place an image close to its Fourier transforming lens. The Fourier transform is unaltered except for the addition of a quadratic (spherical) phase function. If a standard, Vander Lugt correlator is employed, then the same optical set up is used to both record and replay the holograms. Thus, the phase conjugation process, that is inherent in this situation, removes the additional spherical phase front. In the BSO correlator however, different lenses and wavelengths may be used to read and write the MSFs, and the removal of the spherical phase term is not guaranteed.

3.1.1 The Location of the Correlation Plane

In this section, an expression for the location of the correlation plane is derived. It is shown that this location depends on the wavelengths used for reading and writing and on the positions of the input images relative to the lenses. In the previous chapter, expressions were derived for the angles of rays in the Fourier plane. These angles were shown to depend on: their point of origin in the image; the distance of the image from the lens; and the spatial frequency with which they are associated. It is shown here that unwanted, complicated spatial filtering effects can be introduced if the images are not in the front focal plane of the lens.

To determine the location of the correlation plane, three independent sets of optics must be considered. These are: the read-side, the write-side and the correlation plane lenses. This is summarised in figure 3.1. Note that, in this case, it is assumed that the correlation plane is formed by the specific inclusion of a lens after the hologram. It

has been suggested [3.1], that more compact systems can be formed by removing this lens and using a converging reference beam to form the MSF. However, as discussed in the previous chapter, it is undesirable to introduce an angular spread into the holographic reference beam.

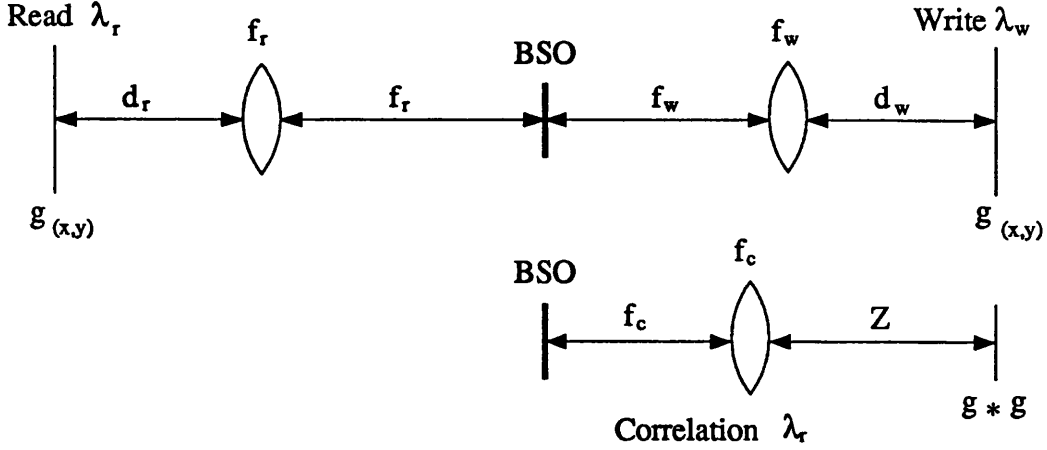


Figure 3.1: Schematic diagram of reading and writing arrangements and the correlation plane.

In the system of figure 3.1, an auto-correlation of the image $g(x,y)$ is being formed. Each replica of the image may be at an arbitrary distance (d_r and d_w) from the two Fourier transforming lenses of focal lengths f_r and f_w respectively. In addition, the read and write-beams may be of different wavelengths; λ_r and λ_w . The Fourier transform of g is multiplied by its conjugate via the hologram formed in the BSO crystal. The result of this multiplication is then Fourier transformed by a lens of focal length f_c to form the correlation plane, which is imaged at a distance Z behind the lens. In order to derive an expression for Z , two approaches will be used. One of these involves an examination of the mathematical expression for the optical Fourier transform, and the second involves a simple geometrical optics approach.

Referring to figure 3.1, the amplitude distribution at the BSO, formed by the transform of g on the read-side is (from [1.5] for example) :

$$A = G_{(u,v)} \exp \left(i \left(x_F^2 + y_F^2 \right) \left(1 - \frac{d_r}{f_r} \right) \frac{k_r}{2f_r} \right) \quad (3.1)$$

Where G denotes the Fourier transform of g , $u = x_F/f_r\lambda_r$ links the Fourier plane coordinates to frequency space and $k = 2\pi/\lambda_r$. If the image g is not in the front focal

plane of the lens – i.e. $d_r \neq f_r$ then the quadratic phase factor, described by the complex exponential term, is non-zero. This is equivalent to the Fourier transform being superimposed on a spherical wavefront.

The spherical wavefront effect can be easily demonstrated by considering an input image $g(x, y) = \delta(x, y)$, where δ is the Dirac delta function, i.e. a point source. If this source is placed in the front focal plane of the lens then a plane wave results; the exact Fourier transform of the delta function being unity over the Fourier plane. However, if the delta function is moved away from the lens, the beam is converging through the Fourier plane.

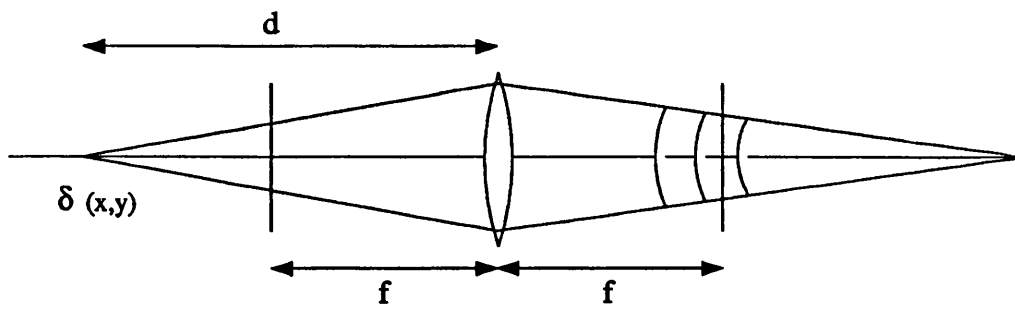


Figure 3.2: Fourier transforming a Dirac delta function

In terms of geometric optics, the lens is forming a real image of the point source, as illustrated in figure 3.2. Conversely, if the delta function is moved towards the lens then a diverging beam is formed. In this second case, the lens forms a virtual image of the point source. It may be noted from equation 3.1, that if $d_r > f_r$ then the argument of the exponential is negative. Thus, it can be stated that, if the argument of the exponential is negative; the spherical wavefronts produced are converging onto a point to the right of the Fourier plane. Conversely, if the argument of the exponential is positive, $d_r < f_r$, then the spherical wave is diverging from a point to the left of the Fourier plane.

Now consider the case when the Fourier transform of a Dirac function is recorded in the BSO hologram. This is equivalent to recording a simple holographic lens. To record the lens, the function $\delta(x, y)$ is placed at a distance d_w from the lens f_w of

figure 3.1, i.e. on the write-side of the system. The amplitude recorded is:

$$A_{hol} = 1 \times \exp \left(i \left(x_F^2 + y_F^2 \right) \left(1 - \frac{d_w}{f_w} \right) \frac{k_w}{2f_w} \right) \quad (3.2)$$

To read this hologram with a plane wave a second delta function is placed in the front focal plane of lens f_r . The output of the hologram at the hologram plane, is then $A_r \times A_{hol}^*$, where $*$ denotes the complex conjugate. Hence the result is:

$$A_{out} = 1 \times 1 \times \exp \left([-i] \left(x_F^2 + y_F^2 \right) \left(1 - \frac{d_w}{f_w} \right) \frac{k_w}{2f_w} \right) \quad (3.3)$$

If $d_w > f_w$ then the argument of the exponential in equation 3.2 is positive, i.e. the diffracted read-beam waves are diverging. This is consistent with the geometric optics/phase conjugation picture shown in figure 3.3.

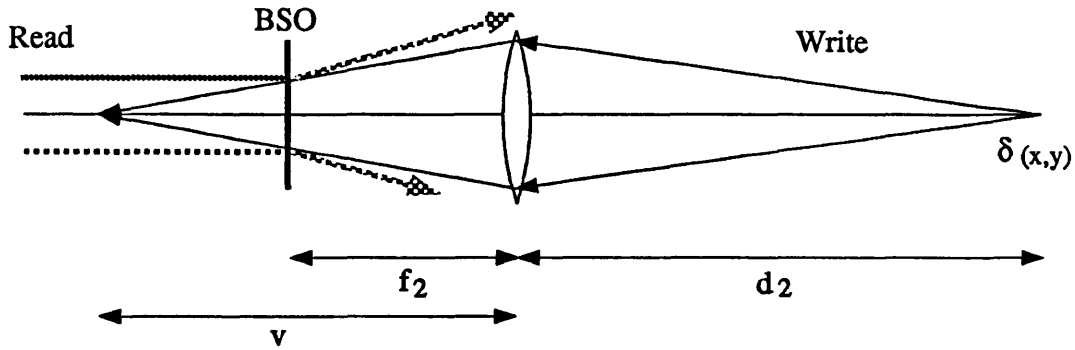


Figure 3.3: Simple holographic lens replayed by plane wave

Now consider the case in which this holographic lens is read out by a spherical wave formed by placing a point source at $d_r \neq f_r$. The amplitude immediately after hologram is then given by:

$$A_{out} = \exp \left(i \left(x_F^2 + y_F^2 \right) \left(1 - \frac{d_r}{f_r} \right) \frac{k_r}{2f_r} \right) \exp \left([-i] \left(x_F^2 + y_F^2 \right) \left(1 - \frac{d_w}{f_w} \right) \frac{k_w}{2f_w} \right) \quad (3.4)$$

This result may be written as $A_{out} = \exp(i\pi C(x_F^2 + y_F^2))$ where C is given by:

$$C = \frac{1}{\lambda_r f_r} + \frac{d_w}{\lambda_w f_w^2} - \frac{d_r}{\lambda_r f_r^2} - \frac{1}{\lambda_w f_w} \quad (3.5)$$

If C is positive then the beam is diverging after the hologram, if C is negative then it is converging.

This approach can again be compared with a simple model based on geometric optics. The lens formed in the hologram may be assessed with reference to figure 3.3. The focal length of the holographic lens, viewed from the write-side of the system, is the distance between the hologram and the point at which a plane wave reconstructing it would either appear to, or actually focus (the lens may be negative or positive). This is simply given by:

$$f_h = v - f_w = \left(\frac{f_w d_w}{d_w - f_w} \right) - f_w = \frac{f_w^2}{d_w - f_w} \quad (3.6)$$

However, light from the read-side of the system sees a lens of opposite sign, the strength of which is altered by a factor of λ_w/λ_r . Thus:

$$f_h = \frac{f_w^2}{f_w - d_w} \left(\frac{\lambda_w}{\lambda_r} \right) \quad (3.7)$$

The input on the read-side may again be considered as a point source, this time being imaged by a two lens system. This system comprises lenses of focal lengths f_r and f_h , separated by a distance f_r as shown in figure 3.4. The final image of the point source is at a distance v_{hol} from the hologram. If v_{hol} is positive, then the rays converge after the hologram; if it is negative, then they diverge.

The first lens forms an image of the point source at a distance v_r from the lens. This image becomes the object for the second lens, the lens to object distance being $f_r - v_r$, i.e. the lens separation minus v_r .

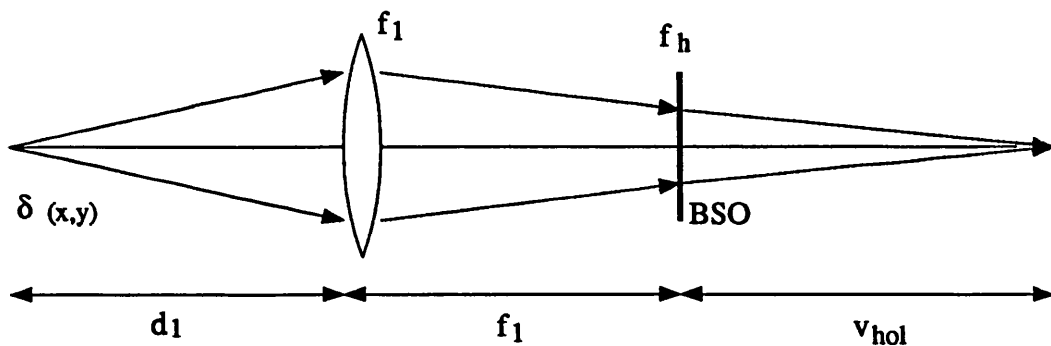


Figure 3.4: System formed of lens f_r and holographic lens

It may be shown that v_{hol} is given by:

$$v_{hol} = \frac{f_h f_r^2}{f_r^2 + d_r f_h - f_r f_h} \quad (3.8)$$

Substituting for f_h from equation 3.7 it is found that:

$$v_{hol} = \frac{\lambda_w f_r^2 f_w^2}{\lambda_r f_r^2 f_w + \lambda_w d_r f_w^2 - \lambda_r f_r^2 d_w - \lambda_w f_r f_w^2} \quad (3.9)$$

From equation 3.9 it can be shown that in order for v_{hol} to be negative:

$$\frac{d_w}{\lambda_w f_w^2} + \frac{1}{\lambda_r f_r} > \frac{d_r}{\lambda_r f_r^2} + \frac{1}{\lambda_w f_w} \quad (3.10)$$

If v_{hol} is negative then the wave fronts are diverging after the hologram and, from equation 3.5, it is seen that this is the same requirement as that for the argument of the quadratic phase factor to be positive. Hence the two methods again produce consistent results.

It should be remembered that the correlation plane is formed by a lens of focal length f_c , located at a distance f_c from the hologram. To derive the required expression for Z , the inverse problem may be considered. That is: at what distance, from lens f_c , should a transparency of the correlation plane be placed, such that its transform would match that calculated by multiplying the transforms of the write and read-images.

In order to produce a solution, the images considered are again Dirac delta functions, the auto-correlation of which is another Dirac delta function. Since the phase factor is radially symmetric, the comparison can be carried out along the x_F axis without loss of generality. The problem may then be expressed as the need to satisfy the following equation:

$$\begin{aligned} \exp \left([-i] x_F^2 \left(1 - \frac{Z}{f_c} \right) \frac{k_r}{2f_c} \right) &= \exp \left(i x_F^2 \left(1 - \frac{d_r}{f_r} \right) \frac{k_r}{2f_r} \right. \\ &\quad \left. + [-i] x_F^2 \left(1 - \frac{d_w}{f_w} \right) \frac{k_w}{2f_w} \right) \end{aligned} \quad (3.11)$$

where the hologram output, given by equation 3.4, is compared with the conjugate of the Fourier transform of the correlation plane. By comparing the arguments of the

exponentials, it can be shown that the value of Z satisfying this condition is given by:

$$Z = f_c^2 \left[\left(\frac{1}{f_c} + \frac{1}{f_r} - \frac{d_r}{f_r^2} \right) + \frac{\lambda_r}{\lambda_w} \left(\frac{d_w}{f_w^2} - \frac{1}{f_w} \right) \right] \quad (3.12)$$

It is known that the required correlation function, for two input Dirac delta functions, is another Dirac function. The input point source on the read-side can be considered as being imaged, by a three-lens system, to form the correlation plane. These three lenses are: the lens of focal length f_r ; the holographic lens; and the lens of focal length f_c . The image distance, corresponding to the two lens system, f_r and the holographic lens v_{hol} , has already been derived in equation 3.9. This image now becomes the object for lens f_c , the distance between object and lens being $f_c - v_{hol}$. The final lens to image distance Z , is the distance from lens f_c to the correlation plane. This is obtained from:

$$\frac{1}{f_c - v} + \frac{1}{Z} = \frac{1}{f_c} \quad (3.13)$$

On substituting for v_{hol} , an identical expression to that given in equation 3.12 is obtained.

Another test case was also considered, in which an image $g(x, y)$, at a distance d_r from lens f_r , is correlated with a Dirac delta function in the front focal plane of lens f_w . The result of this correlation should be a re-imaged version of $g(x, y)$. This is obvious since the input image is transformed, diffracted from a plane grating and re-imaged by lens f_c . On making the substitution $d_w = f_w$ in equation 3.12 it is found that:

$$Z = f_c + \frac{f_c^2}{f_r} - \left(\frac{f_c}{f_r} \right)^2 d_r \quad (3.14)$$

This is simply the equation for the image distance of a telescope formed by two lenses of focal lengths f_r and f_c , separated by a distance $f_r + f_c$, which is again the correct result.

3.1.2 Unwanted Spatial Filtering Effects

After this analysis had been carried out, a second effect due to the position of the input images was noted. In this case, movement of the input plane away from the

front focal plane of the lens alters the angular spread of the light in the Fourier plane. This is seen from equation 2.29, of chapter 2, when $D \neq f$. In this case the beam angles at the hologram are a function of both the position in the input image and the spatial frequency.

To illustrate these combined effects, the expression for the beam angles was combined with equation 2.1 of chapter 2. Equation 2.29 allowed the calculation of a ray angle, produced by any combination of image position and spatial frequency. Equation 2.1 was then used to predict the diffraction efficiencies, associated with the various angles. The case of a $10\mu\text{m}$ grating, in a 1mm thick hologram, probed by a 633nm read-beam was used. Rays originating from three positions within the input plane were considered. The distances of these positions from the optical axis were expressed as a fraction of the focal length of the lens, f : they were 0.0, $0.01f$ and $0.02f$. For each position, the diffraction efficiency was plotted as a function of spatial frequency. It should be noted that the spatial frequencies were either negative or positive depending on which half of the Fourier plane was being considered.

Two graphs are plotted for the read-beam: the first, figure 3.5, illustrates the case when the input plane is at a distance of $0.75f$ from the lens; the second, figure 3.6, illustrates the case when this distance is $0.25f$. In plotting the graphs, account was taken of the change in angle as the rays enter the BSO. It was assumed that the optical axis of the Fourier transforming lens lay at the Bragg angle of the hologram. Therefore, at the zero spatial frequency point, the efficiencies for the three different rays show the angular selectivity effects across the input plane. For both input plane to lens distances, a pronounced, asymmetric variation of the diffraction efficiency was observed as a function of spatial frequency. This would correspond to a spatial filtering operation. The exact form of the filtering depends on the point from which the rays originate in the input plane; i.e. the image would be subject to position variant filtering. This in turn implies that the effects could not be compensated for by any sort of global filtering of the image's Fourier transform, prior to the BSO. Thus, if it is wished to avoid complicated spatial filtering effects, the input plane for the read-image should be located close to the front focal plane of the Fourier

transforming lens.

Figure 3.7 shows the same effects computed for the writing object beam. This plot is for a distance, between input plane and lens, of $0.25f$ but shows a negligible variation with spatial frequency. This is because of the much smaller Bragg -tuning caused by changes in write-beam angle as discussed in chapter 2.

Although an expression has been derived for the location of the correlation plane, its position in the actual optical system will be measured by carrying out test correlations. If an edge enhanced auto-correlation is performed, the output detector can be moved through the correlation plane in order to locate the peak signal. In theory, the reference beam should be balanced to the high spatial frequencies which should produce the largest through-plane variation in the correlation output. This is analogous to the through focus change when an apertured, collimated beam is focussed down by a lens.

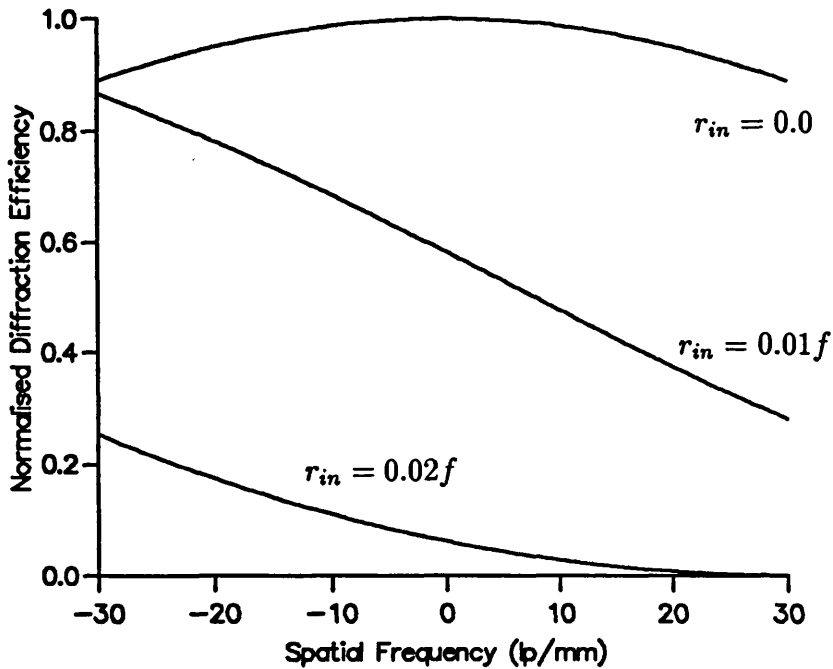


Figure 3.5: Variation of diffraction efficiency with read-image spatial frequency for points at 0, $0.01f$ and $0.02f$ from the optic axis. The input plane is $0.75f$ from the Fourier transforming lens.

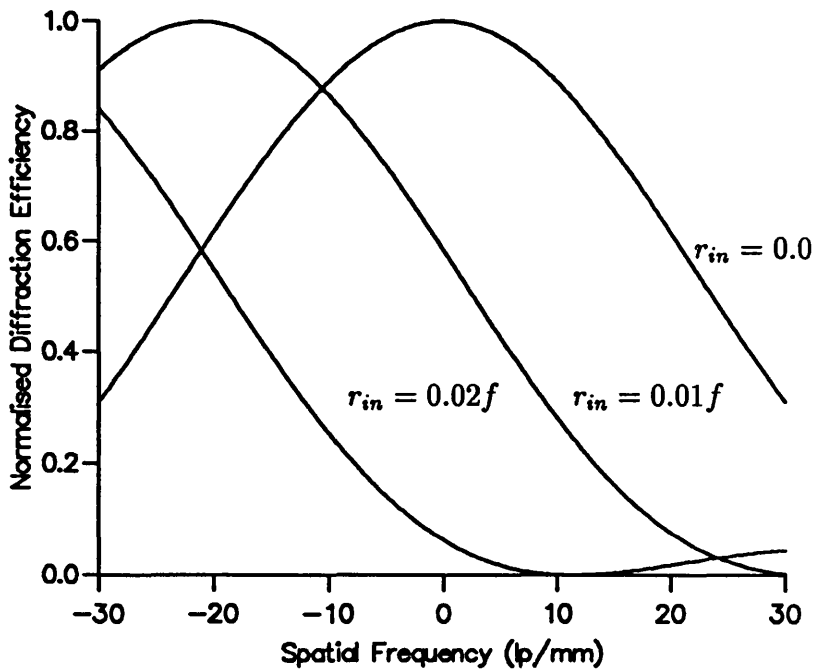


Figure 3.6: Variation of diffraction efficiency with read-image spatial frequency for points at 0, 0.01 and 0.02 f from the optic axis. The input plane is 0.25 f from the Fourier transforming lens.

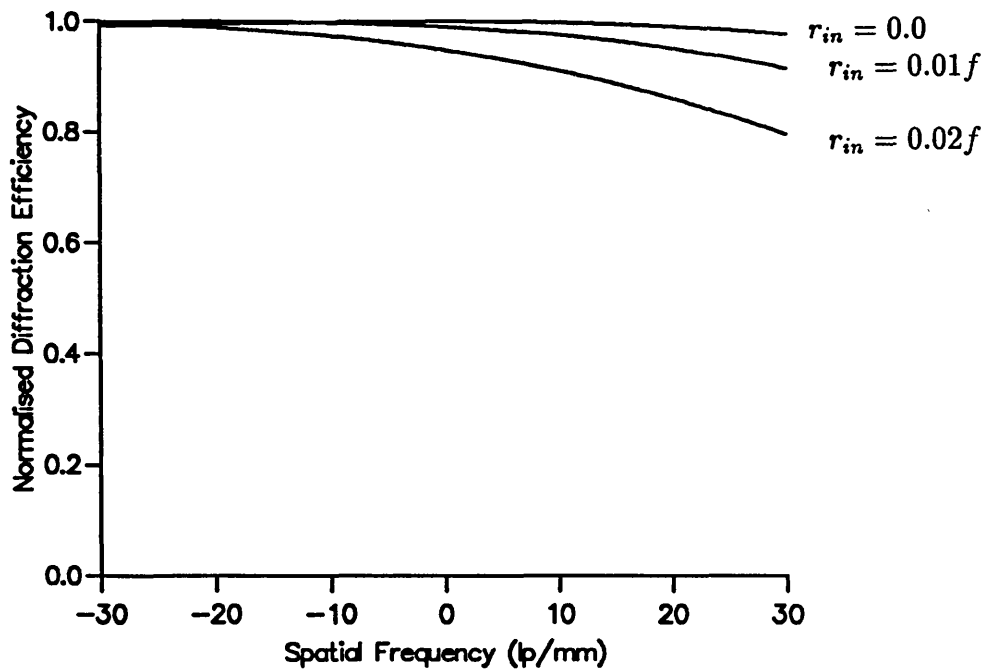


Figure 3.7: Variation of diffraction efficiency with write-image spatial frequency for points at 0, 0.01 and 0.02 f from the optic axis. The input plane is 0.25 f from the Fourier transforming lens.

In this case, the larger the aperture, the more rapid is the expansion of the beam around the focal point. By balancing the reference beam to the higher frequencies of the MSF, light from the edges of the Fourier plane is being collected by lens f_c which focusses it down to form the correlation peak.

In order to test this hypothesis, a computer program was written to produce defocussed correlation planes. The method of achieving this was to form an auto-correlation, with a spatial frequency weighting to simulate the holographic storage characteristics of the BSO; this is discussed in more detail in a later section. The auto-correlation result was stored as a data file of complex numbers. A second program was then used to inverse-transform this file, in order to return to the Fourier plane, and then impose a spherical phase front onto the transform. When this was re-transformed, the result was a defocussed correlation plane. At the time of writing, the amount of defocussing was only controlled by the curvature of the spherical wavefront applied to the transform; so the results were not related to actual lens, image or wavelength parameters. However, it was only necessary to demonstrate the fact that balancing the reference beam amplitude to higher spatial frequencies would increase the through focus variation. Thus, actual numbers were not required.

The test image used was a 28×28 pixel square embedded in a 64×64 pixel image. Results are shown in figure 3.8, for a reference beam with an amplitude 4 times lower than that of the d.c. of the Fourier transform. The figure shows a slice through the maxima of three auto-correlation peaks. On the left is the 'perfect' result, while in the centre is the result of adding a spherical phase front with a maximum difference of π radians between the centre and the edge of the Fourier plane. On the right of the figure is the same correlation result, but with the distortion increased to 2π radians; increasing the defocussing. A repeat experiment with an amplitude 32 times lower than that of the d.c. is shown in figure 3.9. The much bigger change in the correlation peak intensity would obviously ease the task of locating the true correlation plane.

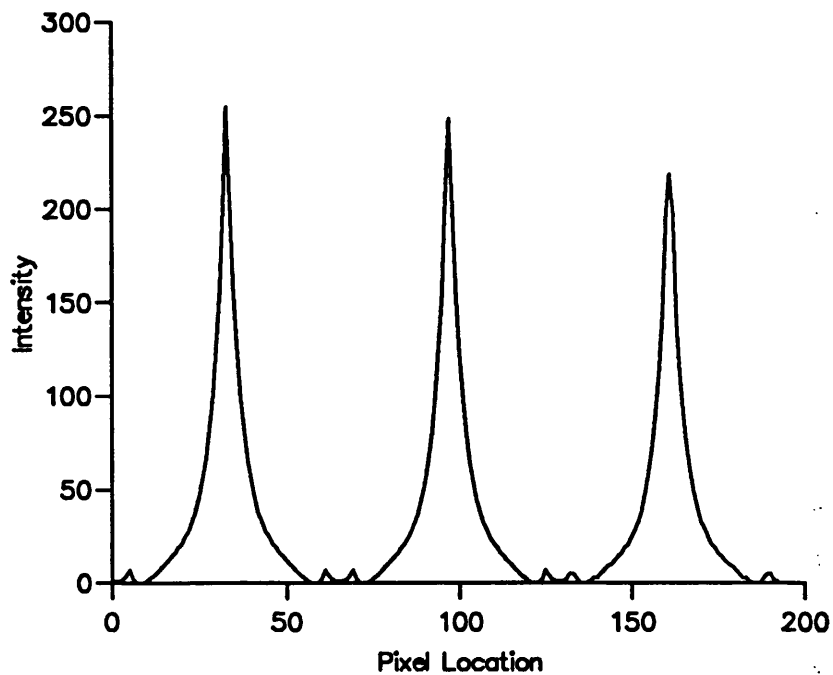


Figure 3.8: Slices through three computer generated correlation planes with a 'perfect' result on the left and progressive degrees of defocussing in the two other peaks. The BSO hologram was simulated with a reference/d.c. beam ratio of 1:4.

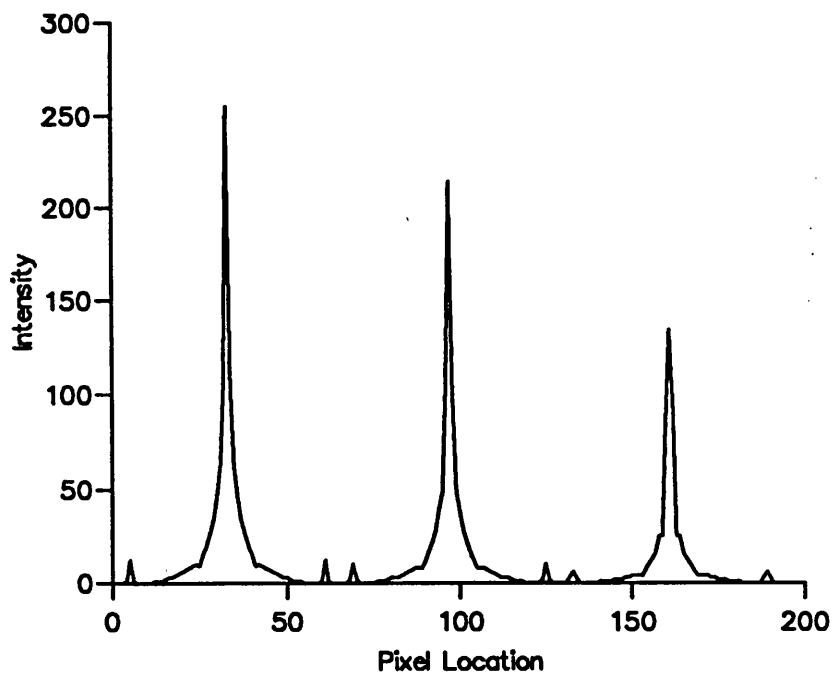


Figure 3.9: Repeat of the above but with a reference/d.c. beam ratio of 1:32.

3.2 Lens and Wavelength Requirements

It has already been shown that the angular selectivity of the hologram in the Fourier plane of the correlator can reduce diffraction efficiency and restrict the location of the input images. It is imperative to reduce, as far as possible, the angular spread of the light originating from the read and write-images. One important factor to consider is the lens' focal length, which should be maximised to reduce the angles in the Fourier plane; as can be seen from equation 2.29 in the previous section. However, there are limits to the focal lengths that can be used. This is because of the physical size of the Fourier transform which is equal to $2f\lambda u_{max}$; where u_{max} is the maximum spatial frequency in the image. This size should not exceed the size of the BSO crystal available for recording the hologram. If it does, vignetting will take place and spatial frequency information will be lost. This is equivalent to a reduction in the resolution of the system. If the limits of the lens' focal length have been reached, and the hologram is still too angularly selective then one solution is to use a larger aperture BSO crystal. This will allow a further increase in focal length, reducing the angular spread. An increase in the size of the crystal however, may not be straight forward, because of the difficulty of maintaining a uniform applied field. The only other alternatives are to use a thinner crystal, at the expense of diffracted energy, or to increase the fringe spacing.

It was decided to use a read-beam wavelength of 633nm. The reason for this was the relative insensitivity of the BSO to light of this wavelength which should reduce the chances of the read-out process perturbing the hologram. This was considered especially important, since, if it could be read out with relatively high intensities, the hologram could be thinner or the applied electric fields smaller. The thinner hologram would be less angularly selective; while applying a smaller field would lead to an increased speed of response. The writing beams would be provided by a cw Ar⁺ laser and would therefore be at a wavelength of 514nm. Since the write and read-images would be relayed to the optical system by identical CRT/SLM combinations, the maximum spatial frequency, u_{max} would be the same for each image. Given a

BSO crystal with a clear aperture of $2r_{BSO}$, it would be necessary to arrange for the products of the focal lengths and wavelengths of the write and read-sides to conform to:

$$f_w \lambda_w = f_r \lambda_r = \frac{r_{BSO}}{u_{max}} \quad (3.15)$$

Thus, different lenses are required for the write and read-sides of the correlator, their focal lengths being determined by the BSO aperture.

3.2.1 Two-Element Compound Lenses

The use of arbitrary focal lengths could be an unattractive proposition, since custom made lenses would be relatively expensive. Furthermore, new lenses would have to be made each time it was desired to change the system parameters. Fortunately, a simple compound lens can be constructed from two 'off-the-shelf' components, to provide a flexible solution to the problem.

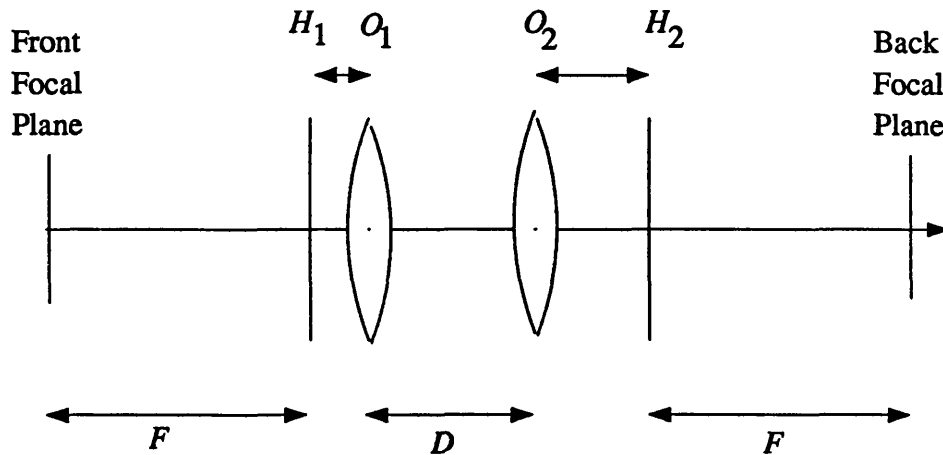


Figure 3.10: A two element compound lens.

Figure 3.10 shows a simple compound lens comprising two 'thin lenses' of focal lengths f_1 and f_2 , at locations O_1 and O_2 , separated by a distance D . The compound lens has a focal length F and principal planes at H_1 and H_2 . The focal length of the compound lens is given in [3.2] as:

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{D}{f_1 f_2} \quad (3.16)$$

The locations of the principal planes relative to the lenses are given by:

$$O_1H_1 = \frac{FD}{f_2} \quad \text{and} \quad O_2H_2 = -\frac{FD}{f_1} \quad (3.17)$$

These lengths are positive if the principal plane lies to the right of the lens and negative if it lies to the left.

One condition that was imposed during lens selection, was that a reasonable 'clear-space' should exist between the individual lenses and the front and back focal planes of the combined element. It was also desirable to select lenses which minimised the overall distance between front and back focal planes. Another important consideration was the required lens diameter. To compute this diameter for both lenses of the combination, ray matrices may again be used. Consider a ray, associated with the maximum spatial frequency in the input image (u_{max}), originating from a point at the edge of the input image which will be designated as $r_{in} = r_{im}$ in equation 2.29. If the input image is situated in the front focal plane of the compound lens then, as can be seen from figure 3.10, this ray will travel a distance of $F - O_1H_1$ before reaching the first lens: remembering that O_1H_1 is negative if the principal plane lies to the left of the lens. After passing through this lens, the ray travels a distance D before reaching the second one of the pair. The ray matrix equations needed to calculate the distance of the ray from the optical axis at each lens are then:

$$\begin{pmatrix} 1 & 0 \\ F - O_1H_1 & 1 \end{pmatrix} \begin{pmatrix} \lambda u_{max} \\ r_{im} \end{pmatrix} \quad (3.18)$$

$$\text{and} \quad \begin{pmatrix} 1 & 0 \\ D & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/f_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ F - O_1H_1 & 1 \end{pmatrix} \begin{pmatrix} \lambda u_{max} \\ r_{im} \end{pmatrix} \quad (3.19)$$

Since the input image was assumed to be square, the required lens radius was $\sqrt{2}$ times the calculated ray-axis distance. Multiplying the matrices and substituting for O_1H_1 gave the following radii for the two lenses:

$$r_{lens1} = \sqrt{2}r_{im} + \sqrt{2}F\lambda u_{max} \left(1 - \frac{D}{f_2}\right) \quad (3.20)$$

$$r_{lens2} = \sqrt{2}F\lambda u_{max} + \sqrt{2}r_{im} \left(1 - \frac{D}{f_1}\right) \quad (3.21)$$

It was noted that $F\lambda_{u_{max}}$ was the radius of the Fourier transform being formed by the lens. Thus, in order to form large Fourier transforms, larger aperture lenses will be needed. Another important parameter was the rate of change of the focal length of the combination with the separation of its elements. If this change is too rapid, it would be difficult to control the focal length of the lens without building the individual components into a sub-assembly. From equation 3.16 it can be seen that the rate of change of F with D is given by:

$$\frac{\partial F}{\partial D} = \frac{F^2}{f_1 f_2} \quad (3.22)$$

Thus, lens combinations with the largest product of individual focal lengths will be the easiest to set up.

It is also be useful to know the errors that might be caused by the assumption that the lenses are thin, and by the errors in the focal lengths of the individual components. The errors quoted in [3.3] for achromatic doublets are $\pm 2\%$ and $\pm 5\%$ for negative singlet lenses. Also, the distance D in equation 3.22 is measured between the second principal plane of the first lens and the first principal plane of the second as shown in figure 3.11.

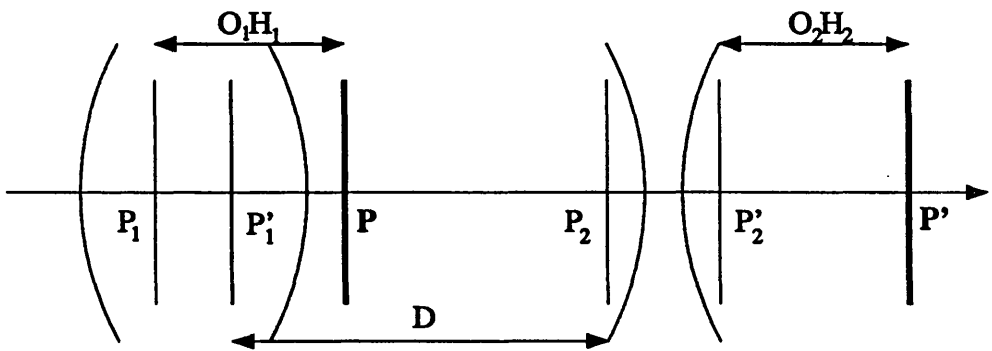


Figure 3.11: Thick lens version of the two-element compound lens.

An inspection of tabulated values in [3.3] suggests that measuring between lens centres will produce typical errors in D of 5mm. This might represent an error of around 5% but it will be different depending on the lens combination chosen. The

error in F can be estimated using:

$$(\delta F)^2 = \left(\frac{\partial F}{\partial f_1} \delta f_1 \right)^2 + \left(\frac{\partial F}{\partial f_2} \delta f_2 \right)^2 + \left(\frac{\partial F}{\partial D} \delta D \right)^2 \quad (3.23)$$

Where δf_1 is the absolute error in f_1 , i.e. $0.02f_1$. The absolute error in D is taken as 5mm. This expression can be evaluated for equation 3.16 to give the absolute error in the focal length. Dividing this expression by F then gives the fractional error. For a system which employs a positive doublet and a negative lens the error is given by:

$$\left(\frac{\delta F}{F} \right)^2 = \left(\frac{0.02(f_2 - F)}{f_2} \right)^2 + \left(\frac{0.05(f_1 - F)}{f_1} \right)^2 + \left(\frac{5.0F}{f_1 f_2} \right)^2 \quad (3.24)$$

where all lengths are in mm. This figure was of relevance in selecting a lens combination to produce a particular focal length.

A simple computer program was written to select, from a set of available lenses, elements for a compound lens of specified focal length. The program calculated the relative positions of the lenses and principal planes, as well as the probable error in the focal length. The rate of change of focal length with lens separation was also calculated. Since an enormous number of combinations were generally possible, restrictions were placed on the selection. These included: limiting the total length between front and back focal planes; stipulating a minimum clearance of 100mm between the lenses and the focal planes; and stipulating a minimum separation of 10mm between the lenses. These restrictions were imposed with a view to facilitating the setting up of the optical system. The lenses were selected from a set including positive focal lengths of 200, 250, 300, 400, 500, and 750mm, and negative focal lengths of -200, -250, -300, -400 and -500mm. The program then ordered the possible combinations according to the total distance between the front and back focal planes.

As an illustrative example, a focal length of 452mm was selected. As will be seen in chapter 5, this corresponded to the expected image resolution of 22.7lpmm^{-1} and a clear aperture in the Fourier plane of 13mm diameter. The wavelength was assumed to be 633nm. In this case, the best combination was found to be a pair of 750mm focal length lenses separated by a distance of 255mm. The total front to back focal plane separation was 851mm, which was less than twice the focal length because the two principal planes are overlapped. The probable error was 1.2% while the value

for $\partial F/\partial D$ was just 0.4, suggesting good tolerance to lens separation accuracy. The required lens diameters would be 27mm for the first lens and 28mm for the second lens, which are also readily obtainable. One of the worst combinations was a positive 200mm lens separated by 89mm from a 200mm negative lens. In this case the total length was 992mm (longer than $2F$), the probable error in focal length was 10.7% and the value for $\partial F/\partial D$ was calculated as -5.1. The required lens diameters were 42mm and 27mm and the first of these is a large diameter for a 200mm lens.

In the above discussion, the criteria for selecting a given lens combination were based on the geometric optical properties of the lenses. However, another important consideration, from the point of view of the correlator, was the effect of lens aberrations on the Fourier transform. Using a very simple extension of the analysis of [1.5], the effect of lens aberrations on an optical Fourier transform can be assessed. In [1.5] a multiplicative phase function was derived to describe a radially symmetric lens comprising spherical surfaces. This expression was inserted into a diffraction integral describing the observed amplitude distribution resulting from the light distribution in a previous plane. In the case of an image placed against the lens, the amplitude at the back focal plane of the lens is given by:

$$U_f(x_f, y_f) = \frac{\exp \left[i\pi(x_f^2 + y_f^2)/(\lambda f) \right]}{i\lambda f} \times \int \int f(x, y) t_l(x, y) \exp \left[\frac{i\pi(x^2 + y^2)}{\lambda f} \right] \exp \left[\frac{-i2\pi(xx_f + yy_f)}{\lambda f} \right] dx dy \quad (3.25)$$

where $t_l(x, y)$ is the lens transfer function. This expression describes a unit amplitude plane wave illuminating a transparency of transmittance $f(x, y)$. This modified amplitude field is then immediately incident on a lens represented by a multiplicative phase function. The light then forms a diffraction pattern U_f at a distance f behind the lens. The lens transfer function, from [1.5], is:

$$t_l(x, y) = \exp[i2\pi n\Delta_0/\lambda] \exp \left[-i\pi(x^2 + y^2)/(\lambda f) \right] \quad (3.26)$$

Note that all expressions were derived assuming that the regions of interest are much smaller than the separation of the planes. The lens transfer function, for instance, when multiplied by a plane wave, produces an approximation to a spherical wavefront for x and $y \ll f$. The first term is merely a uniform phase delay representing the finite thickness of the lens, Δ_0 being its thickness at the optical axis. When the lens transfer function is inserted into the diffraction integral, the terms $\exp[\pm A(x^2 + y^2)]$ cancel. The remaining expression represents the Fourier transform of $f(x, y)$, with respect to spatial frequency components $(u, v) = (x_f/\lambda f, y_f/\lambda f)$. The Fourier transform is not exact, being multiplied by a quadratic phase function, but for correlation this term mainly affects the position of the output plane. In a real system, the ideal approximation, that an incoming plane wave will be converted to a spherical wave, will be inexact. The lens will, in general, introduce aberrations into the transmitted wavefront. Thus, a lens transfer function, including the aberration terms should be substituted into the diffraction integral. In order to do this the nature of the aberrations has to be considered.

The aberrations can be quantified as deviations from perfect performance as illustrated in figure 3.12.

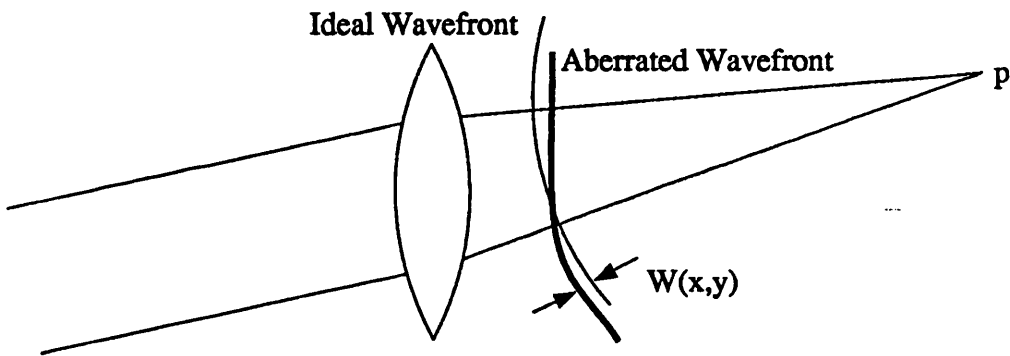


Figure 3.12: Wavefront aberrations produced by a lens.

The collimated beam, entering the lens at some arbitrary angle, should be converted to a spherical wave, ideally converging on the image point p . However, because of lens imperfections, the wavefront immediately behind the lens may take an aberrated form; represented by the heavier line in the figure. All the rays from this wavefront do not travel to the ideal image point, which becomes spread out or blurred. At each

point on the wavefront emerging from the lens, the optical path difference $W(x, y)$ between the actual wave and the ideal spherical wave can be measured. $W(x, y)$ is known as the wavefront aberration and the path difference is measured along the rays from the ideal reference sphere to the image point.

For axially symmetrical optical systems, the wavefront aberrations are shown to depend on ray angles and position of the rays in the exit pupil of the lens ([3.4] for example). However, for the Fourier transforming process, these angles and positions depend on the point in the input image and the spatial frequency with which a given ray is associated. Thus, ignoring a constant phase term, the lens transfer function for a system including aberrations might be written:

$$t_l(x, y) = \exp[iW(x, y, u, v)] \exp \left[-i\pi(x^2 + y^2)/(\lambda f) \right] \quad (3.27)$$

Note that this assumes a small angle approximation since the wavefront aberrations are measured in the z direction rather than along the actual rays. When the expression of equation 3.27 is inserted into equation 3.25 the light distribution in the back focal plane of the lens is found to be:

$$U_f(x_f, y_f) = \frac{\exp \left[i\pi(x_f^2 + y_f^2)/(\lambda f) \right]}{i\lambda f} \times \int \int f(x, y) \exp[iW(x, y, u, v)] \exp[-i2\pi(ux + vy)] dx dy \quad (3.28)$$

Ignoring the quadratic phase term this is seen to be:

$$U_f = \mathcal{F}_{(u,v)} (\exp[iW(x, y, u, v)]) * \mathcal{F}_{(u,v)} (f(x, y)) \quad (3.29)$$

which is the Fourier transform of $f(x, y)$, convolved with the Fourier transform of the wavefront aberration. However, since the wavefront aberration itself depended on the particular spatial frequency being considered, this becomes very difficult to evaluate for functions comprised of many spatial frequencies. The actual position of an object in the input image will also effect the result. Two references were found which dealt with the effect of lens aberrations on the Fourier transforming process [3.5][3.6]. These authors circumvented the problems by only considering images containing a single spatial frequency, that were also centred on the optical axis. The images were

a pupil function filled by a grating, and were placed at the lens. This effectively involved computing or measuring the point spread function of the lens as a function of the angle of tilt applied to it. No attempt was made to extend the analysis to more complex images. The two papers were also concerned with the Fourier transforming process rather than a correlation system.

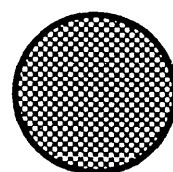
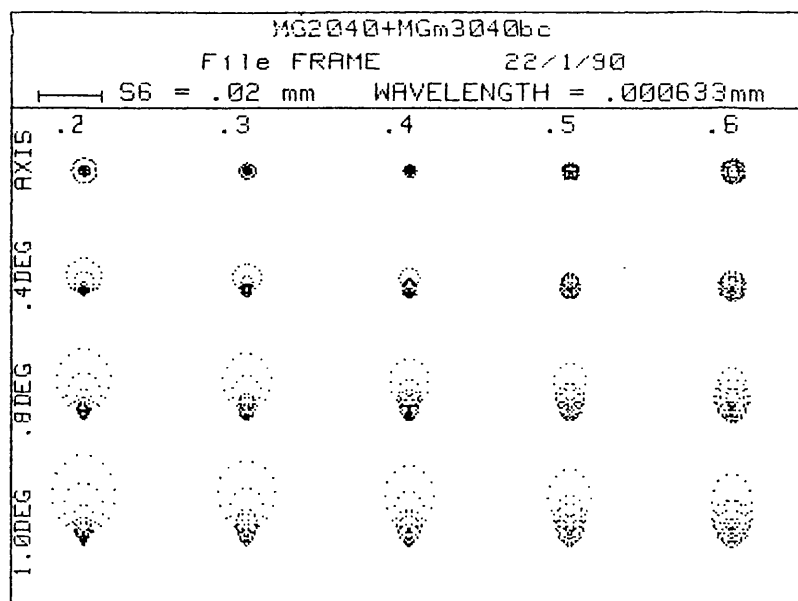
It was noted that [3.5] also considered the effects of aberrations in the collimating optics of the optical transforming system. The wavefront aberration used was of a quadratic form which, although not mentioned in the paper, was equivalent to de-collimating the beam. Thus, the analysis was effectively considering the effects of a defocus in the Fourier plane. Such problems would normally be resolved by carefully setting up the system and ensuring that the BSO hologram was at the observed focal plane of the lens.

In a standard Vander Lugt correlator, any aberration information would be stored, with the reference Fourier transform, in the matched filter. The same lens would be used in the recording and replaying of the filter. If the object were centred for both recording and replaying, it would be possible to form a good auto-correlation as the aberrations would be phase-conjugated by the MSF. The price paid for the presence of the aberrations however, would be a loss of shift invariance in the system. In the case of the BSO correlator, different optics would be used for writing and reading the MSFs. Thus, it may be expected that little correlation would exist between any aberrations. Since their effects would be extremely complex to model, it was decided to choose systems which exhibited negligible aberrations, i.e. the problem would simply be 'designed out'. Aberrations may arise from two sources: firstly, the inherent aberrations of the lenses themselves; and secondly, centring and angular alignment errors. It was assumed that these latter sources of error could be eliminated by careful setting up procedures.

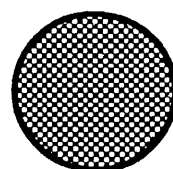
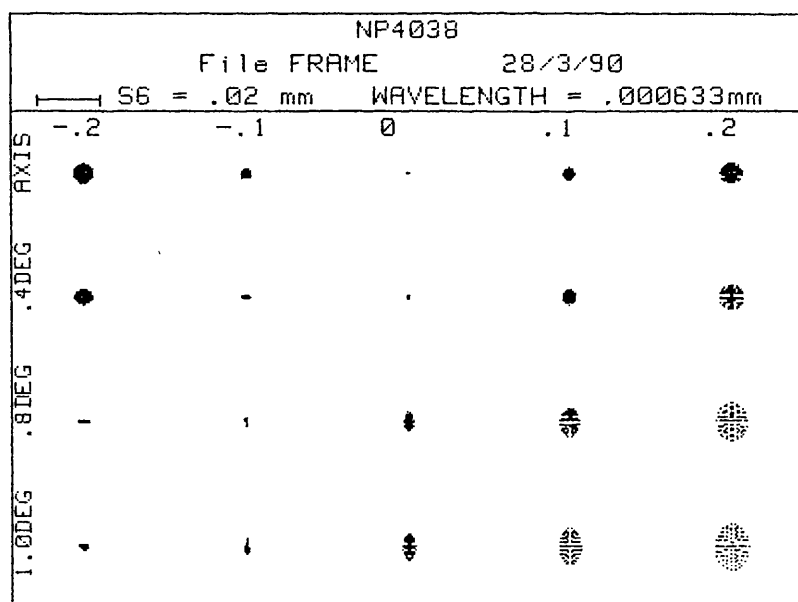
In order to assess the aberrations of the various, two-element lenses, a commercial, ray-tracing program [3.7] was used. The calculation of actual wavefront aberrations was not possible with the program available. However, it was possible to produce geometric point-spread diagrams for the lenses. The computer traced the paths of

rays, through the optical system, from rings of points in a specified input plane. The angles of the rays, relative to the optical axis were also specified. The refraction of these rays at all subsequent optical surfaces was computed and the points where the rays struck the focal plane were plotted. The lens data used by the program, was supplied by the manufacturers Newport and Melles Griot. On examining the data, it was found that, for optimised achromats of a particular focal length, the surface parameters varied with lens diameter. For the negative lenses, on the other hand, the surface parameters were always the same for a given focal length. Thus, while it might be worth selecting the diameter of a doublet to be used in the system, any diameter negative lens would yield the same results.

In an ideal system, all rays of a certain input angle would focus down onto a single point in the Fourier plane. The presence of aberrations however, leads to a spread in the output points. An example of a typical computer output is shown in figure 3.13, with a representation of the diffraction limited point spread function for comparison. The size of the latter was calculated from the diameter of the 'Airy' disk formed by an aperture of radius r_{im} , i.e. $\text{dia.}(\text{psf}) = 1.22\lambda f / r_{im}$. The figure shows results for a Melles Griot achromatic doublet, of 200mm focal length and 40mm diameter, followed by a -300mm focal length, biconcave lens from the same manufacturer. The lenses were spaced so as to form a compound lens of 452mm focal length. The input aperture was situated at the front focal plane of the lens. For comparison, the results for a single Melles Griot achromatic doublet of 400mm focal length and 38mm diameter are also shown. It should be noted that the 1 degree beam angle corresponds to a spatial frequency of $0.017/633 \times 10^{-6} = 28 \text{ lpmm}^{-1}$. As in [3.5], only images of discrete spatial frequency content are considered. The dominant aberrations, shown in figure 3.13, are spherical aberration (defocus), coma and field curvature (change in position of best focus). This latter effect is important, since it would be undesirable for the best focus to 'walk out' of the limited hologram thickness. The actual point spread function of the system is a convolution of the geometric, ray-trace spot and the diffraction limited spot.



Diffraction limited spot.



Diffraction limited spot.

Figure 3.13: Aberrated spot diagrams from computer output with schematic representation of the diffraction limited spot size. Top is a 452mm focal length lens formed from a 200mm doublet followed by a -300mm biconcave lens. Bottom is a single 400mm doublet for comparison.

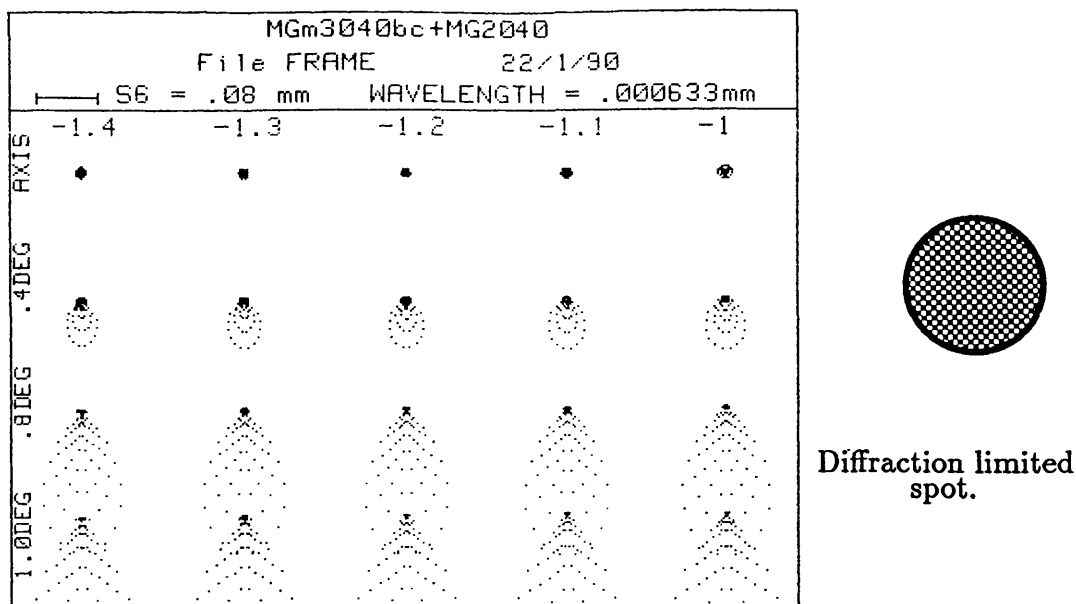


Figure 3.14: Spot diagrams for a 452mm focal length lens in which a -300mm biconcave lens is preceded by a 200mm doublet.

Thus, a criterion for lens selection was a requirement that, up to the spatial frequency of interest, the geometric spread be small compared with the diffraction limited spot size.

Several broad conclusions were drawn from the ray tracing experiments. The results depended, to some extent, on the distance between the aperture and the first lens. If a ray, diffracting from the input aperture, travelled a long way to the lens, then it would strike it at a point near to its edge. If the lens was close to the input plane however, the rays would not get very far from the central region of the lens. The difference in the results for these two cases, arose from the better performance of the lens when the region close to its centre is used. A combination of two positive doublets formed a very well corrected system, superior even to a single doublet of a comparable focal length. On the other hand, any combination in which a negative lens preceded a positive lens showed a very poor performance. An example of this is shown in figure 3.14; in this case, the lenses, used to produce the plots of figure 3.13, have been placed in reverse order, but with their combined

focal length remaining the same. This result is to be expected, since the use of a diverging lens as the first element forces the second lens to 'do more work' in focussing the light down to a point. It was nevertheless possible to select positive/negative lens combinations of sufficiently good performance for the correlator. Since these combinations possessed different spacings between optical elements, this meant that more flexibility in system design was available. From the results, it was clearly seen that the best lens combination, for the formation of a 450mm focal length lens, was a pair of 750mm doublets. This combination actually performed better than a single 400mm doublet, with regard to aberrations. If a 450mm focal length lens were used to transform the read-image, then a 560mm lens would be needed on the write-side. In this case, the best combination was again a pair of 750mm doublets, but the clearance between front and back focal planes and the lenses was quite small, being 194mm. However, other, more than adequate, combinations are available; e.g. a +300mm lens followed by either a -300, -400 or -500. These allow a choice in the distances between the focal planes and the lenses.

It has been pointed out [3.8] that, in many cases, an improved performance can be obtained by a departure from the standard image-lens-transform arrangement. Specifically, the image can be placed just behind a lens, in a converging beam. This has the advantages that no collimating optics are required for the input beam, and a variable, focal length may be achieved with a single lens. More importantly, the lens only needs to be well corrected for rays travelling parallel with the axis, rather than exhibit good performance for a range of input angles. This scheme is ideal for producing Fourier transforms of photographic transparencies. However, it was not a suitable method to use with the polarising beam splitters and SLMs that were to be employed in this project, because of their angular sensitivity.

3.3 Powers Required in the Read and Write-Beams

One advantage of the digital implementation of correlation is the relatively large dynamic range of the digital system. If one set of correlation results include numbers of the order of 10^6 , and other results from the same system are of the order of 10^{-6} ,

there is usually no problem. However, in the optical system, the output correlation plane will be captured by a CCD camera, which usually produces a maximum of 256 usable grey-levels. Therefore, if the correlation plane intensities varied by more than an order of magnitude, the limited dynamic range of the cameras would pose a serious problem. Another important point, is that the absolute values of the correlation results are not important for the digital system; whereas a CCD camera is limited by its sensitivity. Thus, it is important to predict the absolute intensity levels expected in the optical system.

3.3.1 Correlation Plane Intensities

In order to predict the intensities in the correlation plane of an optical system, it is necessary to relate a computed correlation pattern to actual intensity values. Firstly, the Fourier transform of a reference image was computed, and modified to simulate the way in which it would be affected by the holographic recording process. The modified transform was then scaled so that its maximum amplitude equalled the expected peak amplitude transmission of the BSO hologram. Any angular selectivity effects were neglected since they would be overly complicated to model and were expected to be relatively weak on the write-side (see chapter 2). The next stage was to take the transform of the second image, multiply it by the reference and compute the inverse transform to form the correlation plane. The actual intensities could then be calculated from the intensity expected in the input plane while taking account of the difference in size between this and the correlation plane. This point is important since the input images, as displayed on a miniature CRT, were expected to be about twice the size of the detector array of a typical CCD camera. The focal length, f_c , of the lens used to form the correlation plane would therefore be chosen to be half that of the lens used to transform the input image, f_r .

In order to illuminate the SLM, it is usual to expand and spatially filter the beam from a laser using a microscope objective and a pin hole. In order to collimate this beam, a lens is placed a focal length away from the pin hole. The focal length of this lens is important, since it determines the uniformity of the collimated beam's

intensity over the area of the input image. As the Gaussian laser beam expands, after the microscope objective, there will be an intensity profile in the x, y plane (normal to the direction of propagation) given by:

$$I(x, y, z) = I_0(z) \exp(-a(z) \times (x^2 + y^2)) \quad (3.30)$$

where: I_0 is the intensity of the beam at the z axis; $a(z)$ is a constant which determines the beam's Gaussian profile; and z is the distance from the pin hole. The rate of change of both of these quantities is determined by the focussing power of the microscope objective used to expand the beam. At any distance from the microscope objective, there will be some radial distance r_1 , measured from the z axis, at which the intensity falls to $n\%$ of the peak value I_0 (the z dependence will be implicit from now on) i.e.

$$\frac{n}{100} I_0 = I_0 \exp(-ar_1^2) \quad (3.31)$$

Where the identity $r^2 = x^2 + y^2$ has been used. In order to calculate the power contained within a circular area of radius r_1 , the following integral can be evaluated:

$$P(r = 0 \rightarrow r_1) = I_0 \int_0^{r_1} 2\pi r \exp(-ar^2) dr = \frac{\pi I_0}{a} (1 - \exp(-ar_1^2)) \quad (3.32)$$

Assuming that there is negligible light loss at the microscope objective and pin hole, the evaluation of this integral between $r = 0$ and $r = \infty$ is equal to the power of the laser, P_{las} . From the previous equation this implies that $P_{las} = \pi I_0/a$. By substituting for $\exp(-ar_1^2)$ in the previous equation, the fraction of the original laser power contained within the radius r_1 is given by:

$$\frac{P(r = 0 \rightarrow r_1)}{P_{las}} = 1 - \frac{n}{100} \quad (3.33)$$

Thus, if the radius r_1 is the radius at which the intensity has fallen to $n\%$ of the central peak intensity, the area within r_1 contains $(100-n)\%$ of the laser power. This is true for any value of z and it serves to illustrate how inefficient the process of expanding the beam to produce a nearly uniform profile can be. Typically, r_1 might be chosen to ensure that the intensity at the edges of the beam is 95% of the peak, central value. In this case, only 5% of the available laser power is utilised. One

method of improving this is by the use of a pair of holographic ‘beam flattening’ elements to redistribute the light [2.4] but this option was not available for this research.

Since the image is square, and has a half width of r_{im} , the intensity at the SLM could be calculated at this point. However, there is one further consideration necessitated by the possible angular selectivity of the BSO hologram (see chapter 2). In the case of the read-beam, the angular selectivity effects will be removed by using a filter to reduce the intensity at the centre of the image. If the diffraction efficiency for light from the edges of the read-image is η_{min} and that for the centre is η_{max} ; the effect of the transmission filter will be to reduce the apparent image intensity by a factor of η_{min}/η_{max} . Thus, the intensity at the SLM can now be calculated taking all factors into account. This will correspond to the intensity corresponding to the maximum pixel in the read-image; which will be taken to be 255.

The digital Fourier transform process that was used ‘conserved energy’, since the sum of the squares of the pixel amplitude values remained constant after a forward and reverse transform. Thus, once the apparent intensity of the input plane pixels was calculated, the output intensity values could be calculated directly. Thus, if the maximum amplitude-squared value in the computed correlation plane is G_{max} , the peak intensity in the correlation plane of the optical system would be:

$$I_{corr} = \frac{\eta_{min}}{\eta_{max}} \frac{f_r^2}{f_c^2} \frac{G_{max}}{255} \frac{0.05 P_{las}}{2\pi r_{im}^2} \quad (3.34)$$

Here it was assumed that the beam was expanded to cover a circular area of radius $\sqrt{2} \times r_{im}$ and that the beam’s intensity falls to 5% of the maximum at this radius. The factor f_r^2/f_c^2 accounts for the difference in size between the correlation plane and the read-image. It is also assumed that the peak value in the reference Fourier transform has been re-scaled to the the value of η_{max} , given by equation 2.4 of chapter 2.

Two potential sources of error, associated with the conservation of energy assumptions, were investigated. The Fourier transform routine that was used to perform

the calculations was configured so that, for the forward transform:

$$\sum_{n=1}^{N \times N} g^2(N)_{in} = N^2 \sum_{n=1}^{N \times N} g^2(N)_{FT} \quad (3.35)$$

and for the inverse transform:

$$\sum_{n=1}^{N \times N} g^2(N)_{FT} = \frac{1}{N^2} \sum_{n=1}^{N \times N} g^2(N)_{out} \quad (3.36)$$

where $g(N)$ is the amplitude value of a pixel. An experiment was carried out to investigate the accuracy of the relationship expressed in the above two equations. A number of forward and reverse transforms were performed on various images of various sizes. The largest error, of 0.1%, was found in a forward transform of a 512x512 image. It was found that the errors were smaller for smaller images. It was also discovered that an error in a forward transform was cancelled out by an equal and opposite error in the reverse transform. Thus, the error in the conservation of energy through a two Fourier transform process was never greater than 0.007%.

Another assumption made in the calculations was that all the energy in the input plane was present in the digital transform plane. For example, consider the computed Fourier transform of an 8x8 pixel square, in a 256x256 pixel image. The result includes the central lobe of the 2-d *sinc* function as well as the next three lobes. However, if the square was displayed on an SLM, more than just three lobes might be present in the resulting optical transform plane. Thus, by conserving energy in the digital Fourier process, the results might contain a positive error, since energy has been included that lies outside the region of the Fourier plane actually computed. To assess the magnitude of this problem, a simple experiment was performed to compute the total energy contained in various parts of the *sinc* function. This was done by summing the modulus-squared of the pixel values in the appropriate areas of the Fourier plane. It was discovered that 99% of the total energy was contained within the first 12 minima. This is the number of minima that would be present if the transform of a 24 pixel square had been computed. Thus, to a first approximation, if objects of greater than 24 pixels were being transformed, it would be expected that the errors due to the assumptions implicit in the energy normalisation process would be less than 1%.

3.3.2 Fourier Plane Intensities

The intensities produced in the correlation plane, depend on the strength of the electric field that can be applied to the BSO crystal. However, the application of the field slows down the response of the crystal to the incident light field. In order to achieve the desired cycle time for the system, it is therefore necessary to produce sufficient write-beam intensities in the Fourier plane. These intensities depend not only on the image being transformed but also on the physical size of the transform produced by the optical system. This in turn depends on: the size and resolution of the input image; the wavelength of the light; and the focal length of the transform lens. Thus, in order to compute the theoretical intensities, using the digitised images, it was necessary to assign a spatial equivalence to the digital representation of the image:

$$N \times N \text{ pixels} \equiv N\beta \times N\beta \text{ mm} \quad (3.37)$$

Here, N was the number of pixels and β , which was determined by the CRT used to write images to the SLM, was in mm per pixel. Thus, assuming the maximum pixel value in the image was again 255, the energy associated with a pixel with an amplitude g was:

$$E = \frac{g^2}{255} \frac{0.05 P_{las}}{2\pi r_{im}^2} \beta^2 \text{ W pixel}^{-2} \quad (3.38)$$

Since there were the same number of pixels in the computed Fourier transform as in the original image array, the Fourier plane was still apparently $N\beta \times N\beta$ mm. Thus, assuming that the output pixels represented energy in Watts, the apparent output intensity for a pixel of amplitude g was:

$$I_{FT} = E/\beta^2 \text{ Wmm}^{-2}. \quad (3.39)$$

It is now necessary to relate the apparent size of the digital Fourier plane to the expected size of the optically produced transform. To do this, it was noted that the maximum 'spatial' frequency in the digital image was $1/2$ cycles pixel⁻¹. Thus, taking into account the scaling factor of β mm pixel⁻¹, the frequency at the extreme

edge of the Fourier plane was:

$$u_{max} = \frac{1}{2\beta} \text{cycles mm}^{-1} \quad (3.40)$$

The apparent distance between this point and the d.c. of the Fourier transform was $N\beta/2\text{mm}$. For the optical system, the size of the Fourier plane was $r_{FT} = f_w \lambda_w u_{max}$ where f_w and λ_w were the focal length and the wavelength used on the write-side of the correlator. Thus, equating r_{FT} with $N\beta/2$ and u with $1/(2\beta)$, the apparent product of focal length and wavelength for the digital system was given by:

$$(f\lambda)_d = N\beta^2 \quad (3.41)$$

The apparent intensity in the output plane must be multiplied by the scaling factor (apparent area)/(real area) given by $(f\lambda)_d^2/(f\lambda)_o^2$. Thus, remembering that, a factor of $1/N^2$ had to be included to conserve energy in the forward transforming process, the final result for Fourier plane intensity was:

$$I_{FT} = \frac{g_{FT}^2}{255} \frac{0.05 P_{las}}{2\pi r_{im}^2} \frac{\beta^4}{(f_w \lambda_w)^2} \quad (3.42)$$

In order to test this theory, an experiment was carried out with a 16x16 pixel square – of grey-level 255 – embedded in a 256x256 field of zeros. The input intensity was assumed to be 0.5 Wmm^{-2} , the product $(f_w \lambda_w) = 0.3$ and β was assumed to be $0.125 \text{ mm pixel}^{-1}$. For this image, the computed d.c. value was equal to $(16 \times 16 \times \sqrt{(255)})^2 = 1.67 \times 10^7$. The intensity at the d.c. of the optical Fourier transform was then computed to be 88.888 Wmm^{-2} . This agreed exactly with theoretical predictions for the optical transform [1.5]; which, for a square aperture, predicts:

$$I_{FT} = 0.5 \frac{l^4}{(f_w \lambda_w)^2} \text{Wmm}^{-2} \quad (3.43)$$

where: l is the length of the side of the square aperture.

The intensity at the first lobe was computed as 4.31 Wmm^{-2} . This departed from theoretical predictions, with an error of 2.5%. In relation to the correlator design, this error was not a problem. This was because the system was being modelled with the reference beam intensity expressed as a fraction of that of the d.c. of the Fourier transform. Thus, it was only the d.c. intensity calculations that were important.

It must be remembered that, as well as the Fourier transform, the laser on the write-side must also provide a reference beam of the same intensity as the region of the Fourier transform of interest. This beam must extend over the whole BSO crystal. Therefore, with the reference beam balanced to the d.c. of the FT, the laser must supply a reference beam with a power of $I_{FT}2\pi r_{BSO}^2/0.05$. This assumes a square BSO crystal of half-width r_{BSO} , and the factor of 0.05 takes into account the spatial filtering and beam expansion.

3.3.3 Response Time of the BSO

In order to relate the computed intensities to the response time of the BSO, an equation for the temporal behaviour of the space charge field from [1.73] was used. By choosing to look at a specific location on the grating, the expression for the space charge field, E_{SC} , of [1.73] was re-written:

$$E_{SC} \propto 1 - \exp\left(\frac{-t}{\tau_g}\right) \cos(\omega_g t) \quad (3.44)$$

This is basically, an exponentially-decaying rate of growth with a modulation term which causes the field to oscillate as it grows. This oscillation results from interference between the refractive index grating, which actually moves as it forms, and the stationary, cosinusoidal intensity pattern. Also from [1.73] it was found that:

$$\omega_g \propto \frac{1}{\tau_g} = \frac{\alpha}{\tau_g} \quad (3.45)$$

Figure 3.15 shows the time development of E_{SC} for various values of the constant of proportionality, α . As can be seen from figure 3.15; in the absence of any oscillations, the time for the field to reach 90% of its maximum value is $t = 2.7\tau_g$. The effect of the oscillations is, primarily, to cause the field to first reach 90% of its steady state value for much shorter times. However, from the point of view of the optical correlator it is undesirable to have any oscillations in the diffraction efficiency as they would make the correlation results difficult to interpret. For this reason it was decided to take times of several τ_g as the limiting response time of the system.

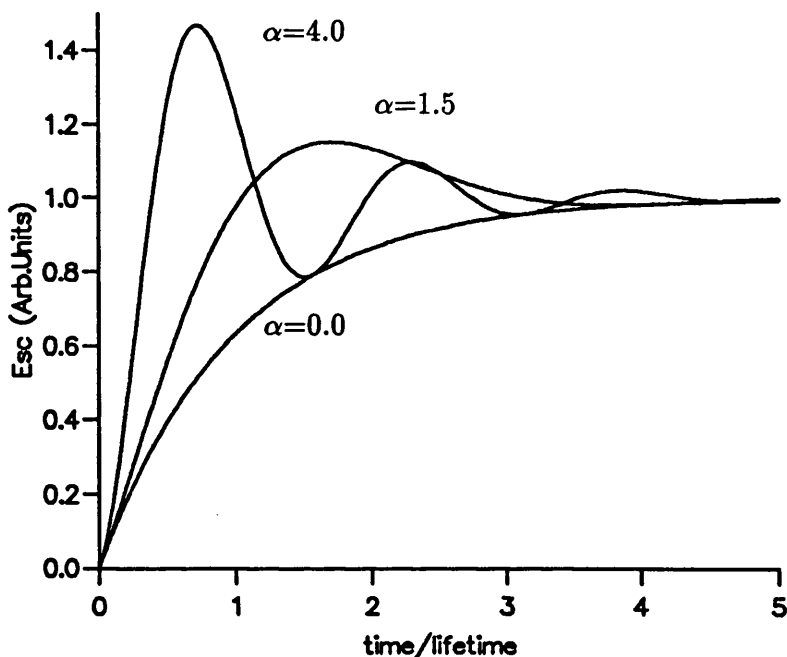


Figure 3.15: The variation of the fundamental component of space charge field with the normalised variable t/τ_g .

The parameter τ_g is given by:

$$\tau_g = \frac{(1 + \Theta^2 \nu^2)^2 + (\Theta^2 \nu E'_A)^2}{(1 + \nu^2)(1 + \Theta^2 \nu^2) + \Theta^2 \nu^2 E'^2_A} \tau_c \quad (3.46)$$

Where:

E'_A is the applied field normalised by the critical field E_C (see chapter 1)

ν is the grating frequency normalised to the critical grating frequency i.e. K/K_C where $K = 2\pi/\Lambda$.

Θ^2 is a diffusion length parameter; being the product of the charge carrier mobility and the free charge carrier lifetime

τ_c is the dielectric lifetime

It is important to note that τ_c is inversely proportional to the intensity of the grating that is forming the hologram in the BSO. Thus, the response time of the BSO correlator will also be inversely proportional to the hologram writing intensity.

In order to predict system response times, the BSO parameters Θ^2 and τ_c had to be measured; the latter at a particular intensity. The oscillatory behaviour of the growth of the space charge field makes measurements based on rise-times difficult. If

decay times are measured under uniform illumination, *in the absence of the writing intensity grating*, the oscillations are not present. However, the decay will still be an exponential function with the time constant τ_g . The most convenient way of measuring the decay rate of the grating was to diffract a probe beam from it and to monitor the decay of the diffracted power with time. The probe beam had to be sufficiently weak, and/or of a wavelength sufficiently removed from the crystals peak sensitivity, so as not to contribute to the grating decay. The diffracted power was proportional to the square of the refractive index modulation of the grating and hence to the square of the space charge field. If the grating decay time, $T_{1/2}$, is the time it takes for the diffracted power to fall to one half of its original value then: $T_{1/2} = \ln(2)\tau_g/2$.

Two experiments, measuring the decay time, were carried out. The first used a fixed grating fringe spacing with various applied fields. The second used a constant applied voltage and investigated the effects of varying fringe spacing. The grating was initially written using an Ar^+ laser and was probed by a weak He-Ne beam. The diffraction was monitored with a photo-diode. One of the holographic writing beams was then shuttered. The grating was erased by the remaining writing beam, and the decay time of the diffracted signal was measured. A preliminary experiment was carried out, to ensure that the expected inverse relationship between decay time and intensity was observed. The results are shown in figure 3.16. Following this, experiments to determine Θ^2 and τ_c were undertaken. The results were compensated for laser power fluctuations, being normalised to an intensity of 3mWcm^{-2} . Figure 3.17 shows experimental results for fringe spacings of $10\mu\text{m}$ and $22\mu\text{m}$. The applied field was normalised by the critical field $E_C = 3.1\text{kVcm}^{-1}$ (see chapter 2). For comparison, theoretical curves for these fringe spacings were generated from equation 3.46. The theoretical plots were produced using a value of 160 for Θ^2 and 5.2ms for τ_c . In the second experiment, the decay time was measured while the fringe spacing was varied, at a constant applied field. The results, for an applied field of 7kVcm^{-1} are shown in figure 3.18 and, in this case, the theoretical curve was generated with $\Theta^2 = 120$ and $\tau_c = 5.0\text{ms}$.

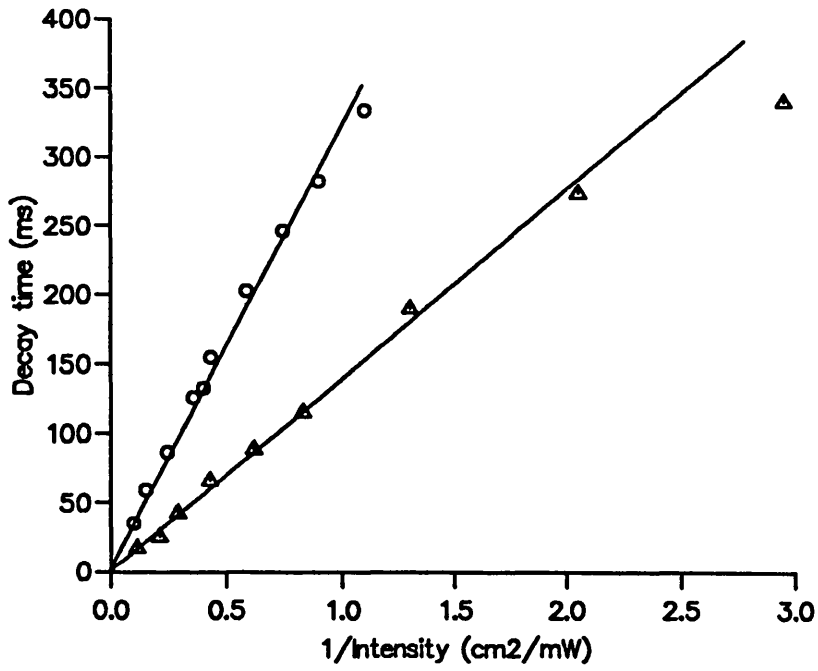


Figure 3.16: Decay time versus erase beam intensity for grating spacings of 10 microns (circles) and 22 microns (triangles).

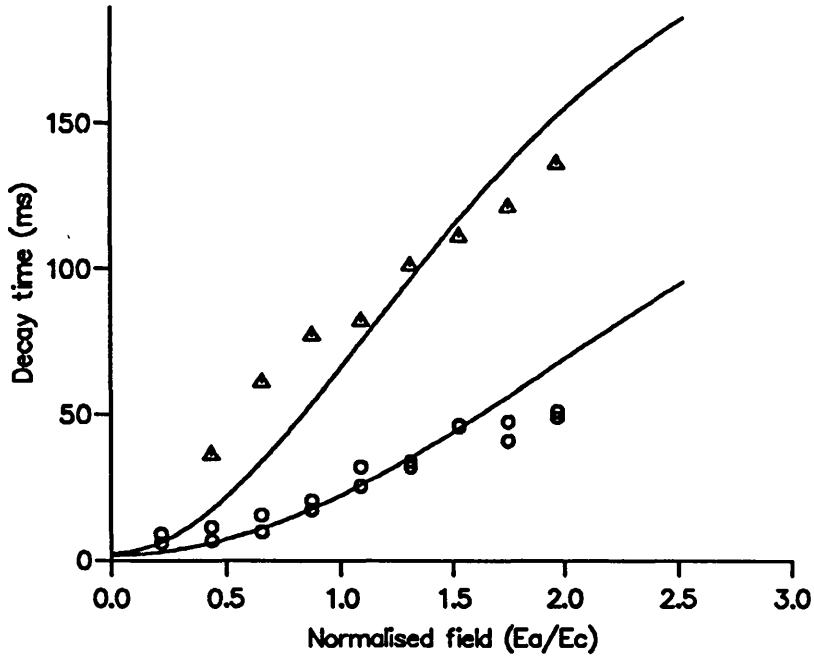


Figure 3.17: Decay time versus, E_A/E_C with grating spacings of 10 microns (triangles) and 22 microns (circles): wash-out beam intensity of 3mWcm^{-2} . Theoretical curves generated with $\Theta^2 = 160$ and $\tau_c = 5.2$.

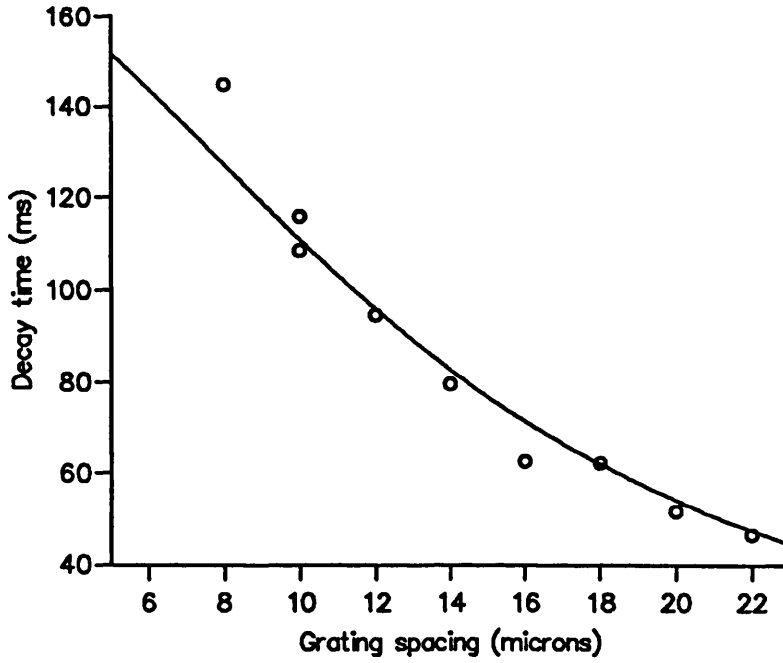


Figure 3.18: Decay time versus grating spacing (circles) for $E_A=7\text{kVcm}^{-1}$: wash-out beam intensity of 3mWcm^{-2} . Theoretical curve generated with $\Theta^2 = 120$ and $\tau_c = 5.0$.

3.4 Tilted Fringes in an Applied Field

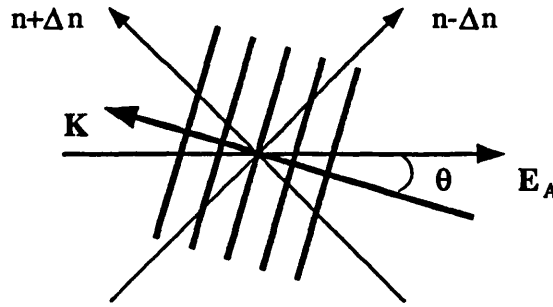


Figure 3.19: Fringes tilted with respect to the applied field.

The optimum geometry, to enhance the holographic grating with an applied electric field, is with the field parallel with the grating vector. One problem with the correlator architecture, is the possibility of the grating vector becoming tilted with respect to the applied field. It will be shown that the possibility of the fringes becoming significantly tilted, increases as the angle between the hologram writing beams is

decreased. The geometry of interest is shown in figure 3.19. As the fringes are tilted, the perpendicular component of the electric field falls.

Some experimental measurements of the effect were made. In order to carry out the experiment, the holographic writing beam directions were held fixed while the BSO crystal was rotated. The polarisations of the writing and read-beams were also rotated in order to keep them at a fixed orientation with respect to the crystallographic axes. This was necessary since these polarisation states may also affect the diffraction efficiency. The results are shown in figure 3.20.

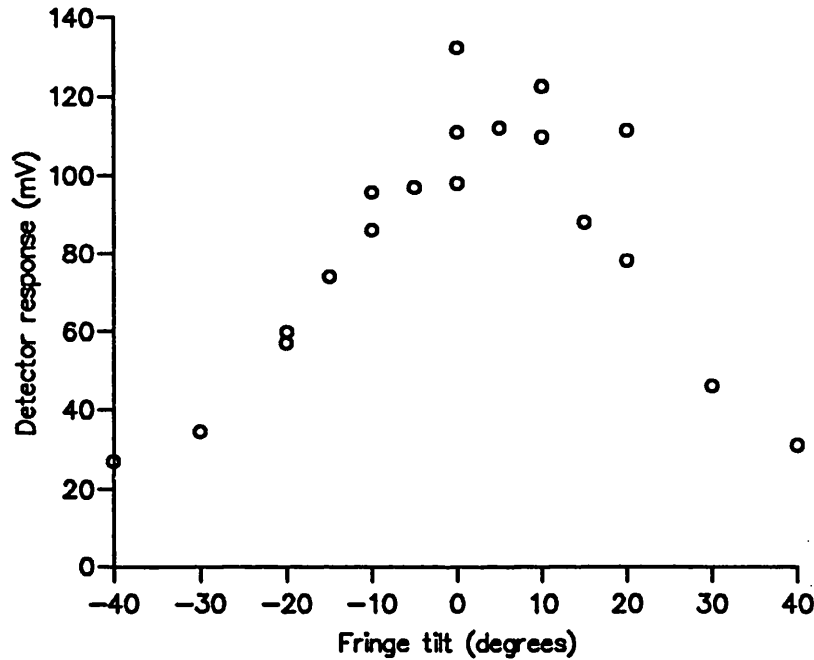


Figure 3.20: Diffraction efficiency versus fringe tilt for a 10 micron grating with an applied field of 7kV. The zero degree point corresponds to a 0 degree angle between grating vector and applied field direction.

A fairly sharp fall off in efficiency was measured as the applied field was tilted with respect to the fringes. The experiment was quite difficult to set up and control; as the scatter of points about the zero degree mark illustrates. The diffraction efficiency, measured as the response of a photo-detector in the diffracted beam, fell to about 50% of its maximum value at tilt angles of $\sim \pm 20^\circ$. A theoretical description of this process is beyond the scope of this thesis. However, it is sufficient, for the correlator design, to note the presence of this effect and to derive an expression for the fringe

tilts that will be produced in the correlator.

In order to predict the fringe tilts that may be encountered, the geometry shown in figure 3.21 was used. The coordinates in the plane of the write-image have their origin at the centre of the image.

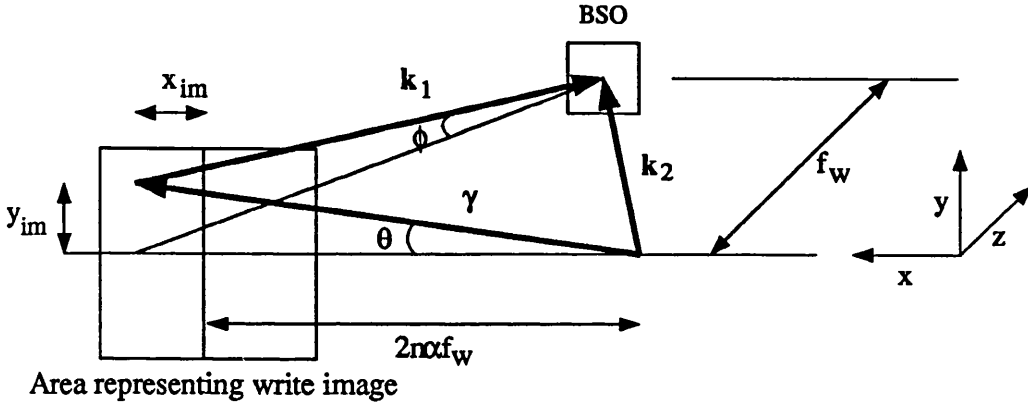


Figure 3.21: Schematic diagram used to determine the fringe tilts produced at the BSO hologram by rays from different parts of the write-image.

This origin was assumed to be located on the optical axis of the Fourier transform lens. The maximum and minimum values are $\pm r_{im}$ at the image borders. Note the directions of the right handed coordinate system in the diagram, with z going into the page. k_1 and k_2 are vectors representing the two beams writing the hologram. k_2 represents the reference beam, and is at a fixed angle with respect to the optical axis of the Fourier transforming lens. The angle of the reference beam is such that it forms fringes of spacing $\Lambda = \lambda_w / 2n\alpha$ when it interferes with light from the centre of the input image. The angle, 2α , is the beam intersection angle in the BSO. The image is assumed to be in the front focal plane of a lens of focal length f_w . Therefore, the angles of the rays in the Fourier plane are determined solely by their point of origin in the input image. When light from the image point x_{im}, y_{im} interferes with the reference beam, the grating vector lies in the plane defined by the two writing beam vectors. This is at an angle θ to the xz plane and hence to the applied field. If β is a unit vector in the x direction, then $\cos(\theta)$ is given by:

$$\cos(\theta) = \frac{\gamma \cdot \beta}{\gamma} \quad (3.47)$$

where γ is the modulus of the vector and:

$$\gamma = \mathbf{k}_2 - \mathbf{k}_1 = (2f_w n \alpha + x_{im})\mathbf{i} + y_{im}\mathbf{j} \quad (3.48)$$

It is then found that:

$$\cos^2(\theta) = \frac{(2f_w n \alpha + x_{im})^2}{(2f_w n \alpha + x_{im})^2 + y_{im}^2} = \frac{(f_w \lambda_w - \Lambda y_{im})^2}{(f_w \lambda_w - \Lambda x_{im})^2 + (\Lambda y_{im})^2} \quad (3.49)$$

where Λ is the fringe spacing formed by light from the centre of the image. From figure 3.21, it can be seen that the worst tilt occurs for image coordinates $x_{im} = -r_{im}$, $y_{im} = \pm r_{im}$.

If a joint transform architecture was constructed, using the same average fringe spacing, then the maximum fringe tilt produced would be even greater. This would be because the reference beam would also subtend a cone of angles. This is another advantage of the Vander Lugt architecture, not previously noted for BSO based correlators. If the hologram becomes thin, the tilt effects can be reduced by using smaller fringe spacings without the penalty of increased angular selectivity.

If the write-image is not in the front focal plane of the lens, then the angles of the rays in the output plane will be a function of spatial frequency. Thus, the write-image could also be subject to position variant spatial filtering. It would therefore be necessary to ensure that the write image be located as close as possible to the lens' front focal plane.

3.5 Limits on Fringe Spacing

In chapter 2 it was shown that, in order to minimise angular selectivity effects, the fringe spacing used should be as large as possible. However, the maximum fringe spacing is limited by the need to ensure the separation of the correlation plane from the undiffracted beam. In [1.5], an expression is derived for this limit. The arguments used, are based on the fact that the pattern on the undiffracted beam

is the auto-correlation of one of the images convolved with the other image. The two images are denoted $h(x, y)$ and $g(x, y)$ and, thus, the on-axis term is given by $h(x, y) * h^*(-x, -y) * g(x, y)$. This is assumed to have a physical width three times that of the original images; as long as the images were initially all the same size. The cross-correlation of h and g on the other hand extends to twice the width of the original images. It is therefore stated that the separation between the centres of these two outputs should be at least 2.5 times the width of the original images. However, it should be noted that this analysis was based on the assumption that the *amplitude* transmittance of the MSF was proportional to the incident interference pattern. In the BSO correlator, phase modulated holograms will be formed. The term giving rise to the triple convolution product should not then be present. Furthermore, the requirement derived in [1.5] may be over stringent if correlation is being carried out in the presence of image edge enhancement. This is because the auto-correlation of a function in the presence of edge enhancement is a sharp peak at the origin. In this case the term $h(x, y) * h^*(-x, -y) * g(x, y)$ will tend to an approximation of $g(x, y)$. The intensities of the undiffracted term may then be negligible outside the area of the original image. The requirement that the centres of the on-axis term and the correlation plane be separated by a distance of $5r_{im}$ was therefore an over-specification. For the purposes of the correlator design carried out here this figure was replaced by $3r_{im}$. This condition could then be linked to the holographic fringe spacing and the read-side lens focal length and wavelength as in [1.5].

Two further rules governing the allowed values of the holographic fringe spacing stem from the predictions of the BTM. Firstly, as the fringe spacing approaches the critical value Λ_C , a regime is reached in which the strength of the space charge field is dependent on the fringe spacing; as shown in figure 2.2. This dependence on fringe spacing is undesirable since there will be a range of fringe spacings produced in the MSF, due to the angular spread in the write-beam. In order to avoid this, the fringe spacing should, if possible be greater than about $6\Lambda_C$. Secondly, faster response times can be achieved with larger fringe spacings.

Unfortunately, the problems that may be associated with fringe tilt can only be

alleviated by increasing the angle between the hologram writing beams. This in turn reduces the fringe spacing and, if the hologram is exhibiting thick behaviour, there will be a trade off between fringe tilt and angular selectivity effects.

Given the need to separate the correlation plane from the undiffracted beam, while keeping the fringe spacing sufficiently larger than Λ_C , the following inequalities must hold:

$$6\Lambda_C < \Lambda < \frac{\lambda_r f_r}{3r_{im}} \quad (3.50)$$

It is interesting to note that this actually imposes a limit on the physical size of the image given by:

$$r_{im(max)} < \frac{\lambda_r f_r}{18\Lambda_C} \quad (3.51)$$

Since the maximum spatial frequency in the image will be limited by the hardware used, the above inequality actually sets a limit on the space bandwidth product of the system. It must be remembered however, that the saturation effects will only be a problem if a significant range of fringe spacings are present in the hologram.

3.6 Comparison of BSO and Holographic Film

The optical Fourier transform of an image, which contains both amplitude and phase information, may be recorded holographically. In early correlation systems, the recording medium was photographic plate. The absolute and relative intensities of the hologram writing beams and the characteristics of the recording medium, will always affect the recorded information in some way. This, in turn, will affect the nature of the correlation function produced. A theoretical study was undertaken in order to compare the storage characteristics of holographic plate with those of the BSO. The effect of nonlinear film response has been studied previously at BAe [3.10], this has been added to here by the creation of a computer simulation, in one-dimension, to allow rapid appraisal of the various effects. In order to describe how storage on film would affect a Fourier transform, a hypothetical curve to describe the variation of amplitude transmission with recording exposure was used. The form of

curve chosen was:

$$T_A = \frac{1}{1 + \delta_1 E^2 + \delta_2 E^3 + \delta_3 E^4} \quad (3.52)$$

Here, E was the incident exposure, which was a product of intensity and time; and the constants, δ , were specified to alter the profile of the curve. This form of curve had the properties of giving a transmission of 1 for zero exposure. It also had a slope that tends to zero as exposure tends to both zero and infinity; which was the reason that no linear term in E was included. A plot of this curve is shown in figure 3.22: with $\delta_1 = 5 \times 10^{-4}$, $\delta_2 = 5 \times 10^{-6}$ and $\delta_3 = 5 \times 10^{-7}$. It can be seen to have a form similar to curves reported in [3.11], for example.

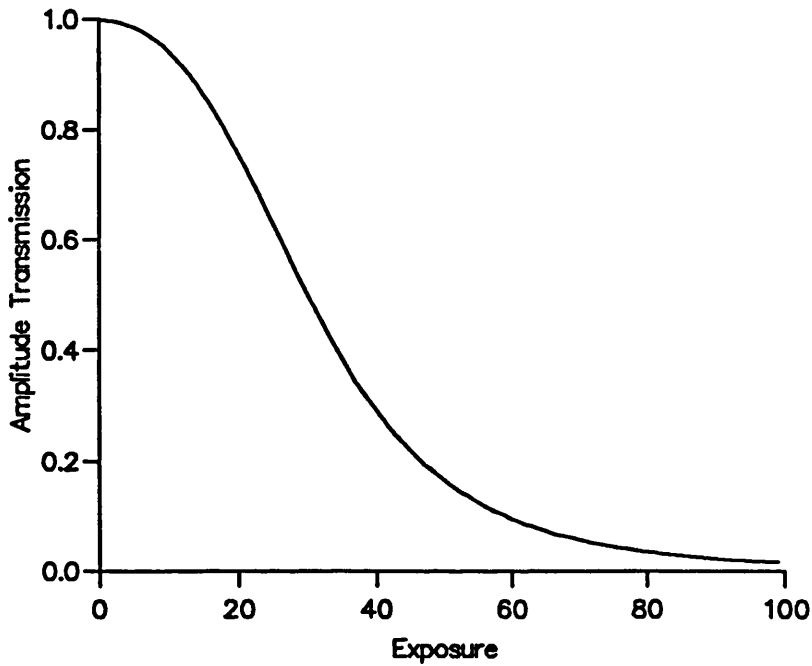


Figure 3.22: Hypothetical film response curve.

In order to assess the effect of this curve on the holographic recording of a matched spatial filter, a 1-D simulation was used. The Fourier transform of a square pulse, 16 points wide, was computed. An ‘interference pattern’ between this transform and a uniform reference beam, of variable intensity, was then formed. To do this, the phase of the transform was preserved but the amplitude was modified according to the relative intensities of the reference beam and each point in the transform. If the intensity of the Fourier transform at some frequency u was $I_{FT}(u)$, and the intensity

of the reference beam was I_{ref} ; then the maximum and minimum intensities in the resulting interference pattern were:

$$I_{int (max/min)}(u) = \frac{1}{2}I_{FT}(u) + \frac{1}{2}I_{ref} \pm \sqrt{I_{FT}(u)I_{ref}} \quad (3.53)$$

Three methods of determining the recorded amplitude of the were investigated for photographic film. The simplest ignored the T_A vs E curve of figure 3.22 and assumed that the amplitude transmission of the film fell linearly with exposure until it reached zero. The amplitude transmittance of the film, after exposure to the interference pattern, was assumed to be sinusoidal, following the incident fringe profiles. The maximum and minimum values of T_A were determined by the supposed linear relationship. The second method was slightly more sophisticated, the amplitude transmittance of the exposed film was again assumed to be sinusoidal but the maximum and minimum values were determined from the T_A vs E curve. Since these first two methods assumed that the recorded fringes were sinusoidal, the amplitude diffraction efficiency at each point in the Fourier transform hologram was calculated from:

$$\eta_{amp}(u) = \frac{1}{4}T_{A \max}(u) - T_{A \min}(u) \quad (3.54)$$

which is taken from [1.5]. The amplitude of the original transform was then set equal to this value.

For the third method, the diffraction efficiency, for each point in the hologram, was computed by performing a 1-D Fourier transform on the modified fringe profile. Typically, the result of this transform no longer contained just the d.c. and the first orders, but also higher orders. The diffraction efficiency was found by measuring the height of the first order peak. To illustrate this process, fringe profiles have been generated for two interfering beams of amplitude 10 (arbitrary units). These gave a maximum intensity of 200 and a minimum intensity of 0 in the resulting interference pattern. One set of fringe profiles was generated by assuming a linear transfer, with the film giving zero transmission at an intensity of 200. The result was an amplitude transmittance varying, as a perfect sinusoid, between 0 and 1. This profile was then modified by assuming that the recording material followed the T_A vs E curve shown

in figure 3.22. The resulting amplitude still varied between 0 and 1, but the effects of saturation were clearly seen in the region of low transmission. The results are shown in figure 3.23. The computed Fourier transforms of the fringes of figure 3.23 are both shown in figure 3.24. Note that the modulus of the amplitude is shown. In order to bring out the detail, the amplitude of the transform of the non-linear fringes was magnified by 3x before plotting.

It was found that the amplitude diffraction efficiency was given by dividing the computed value at the first order by the number of points in the array. This was due to the organisation of the normalisation factors in the FT routine. It was interesting to note that most of the energy in the higher orders had come from the d.c. and the amplitude of the first order was only slightly reduced. Another experiment was carried out in which a beam of amplitude 16 was interfered with a beam of amplitude 4. For linear transfer, sinusoidal fringes were produced, with a T_A varying between 0 and 0.65. However, saturation effects resulted in a T_A that varied between 0 and 0.05, if the nonlinear transfer curve was used. Thus, as expected, saturation effects have attenuated the fringes of low modulation with high intensity.

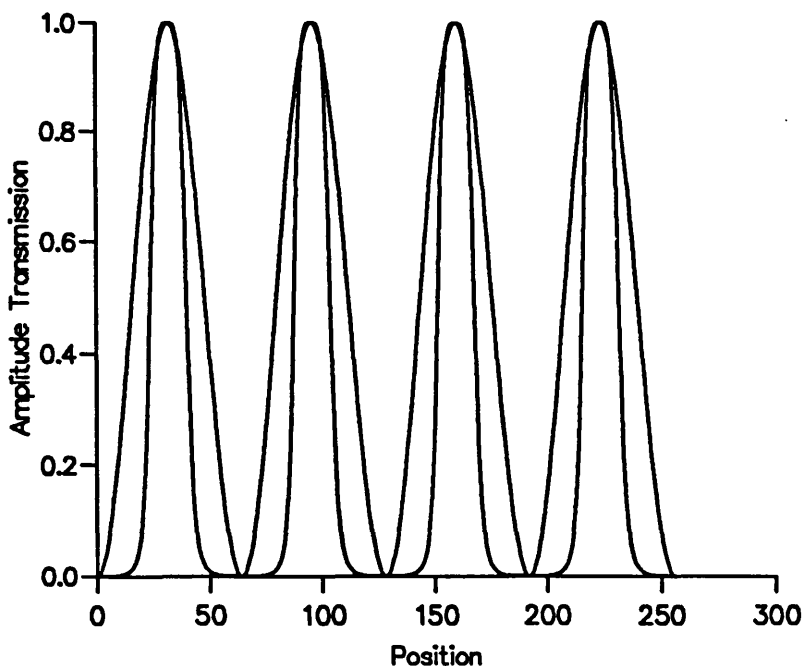


Figure 3.23: Sinusoidal fringes and the result of their non-linear recording.

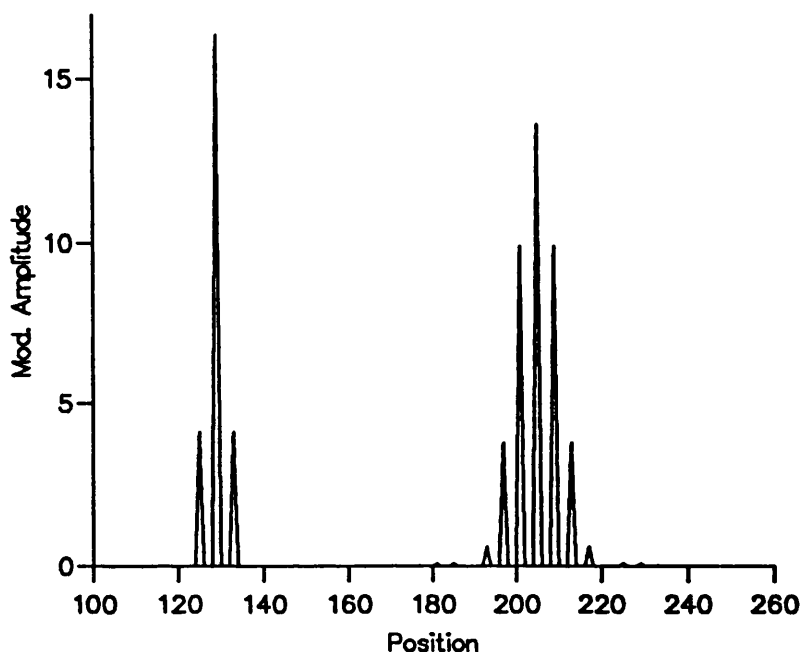


Figure 3.24: Fourier transforms of the perfect sinusoidal fringes (left) and the same fringes subject to the film transfer curve (right).

For the BSO, the space charge field, and hence the amplitude diffraction efficiency, is proportional to m , the reduced fringe modulation. If the cosinusoidal fringe pattern is given by $I_o(1 - M \cos(x))$, then M is equal to $(I_{max} - I_{min})/(I_{max} + I_{min})$; where I_{max} and I_{min} can be obtained from equation 3.53. The reduced modulation m is then given by $m = M/(1 + \beta/(sI_o))$ where β and s are the probabilities for thermal and photo-excitation respectively. Thus, at very low intensities the hologram's modulation will tend to zero. However, in order to simplify matters, it was assumed that $m = M$ and the thermal effects were ignored.

There were now four models to compare: the linear model; the amplitude model; the distorted fringe model; and the BSO model. When the Fourier transform of the 16 pixel square pulse was computed: its d.c. amplitude was 16; the first lobe was -3.5; the second 2.1; and the third -1.6. The first set of experiments involved balancing the reference beam to the d.c. amplitude of the transform.

Initial experiments with the simple linear model suggested that its predictions were not physically realistic. It was therefore decided to concentrate on the amplitude

and the distorted fringe models when studying the response of the photographic film. The amplitudes of the Fourier transform were linearly scaled so as to match the interference pattern intensities to various parts of the film response curve of figure 3.22. The first experiment involved a reference beam of amplitude 16 with the original, unscaled transform. The peak intensity of the fringes was 512; which was well into the curve's saturation region. The result of this was a recorded transform with its amplitude biased heavily towards the d.c. The reason for this was that the d.c. was the only part of the interference pattern where the intensity fell to zero, consequently this was the only place where good modulation was recorded. As the minimum fringe intensity started to rise, the recorded amplitude transmission fell at a much greater rate, since the recording was well into the saturation region of the film. Thus the modulation of the recorded pattern fell much more quickly at points away from the d.c. than it did in the original fringe pattern.

At this point it was noted that care had to be taken in the interpretation of these results. When deep into the saturation region, the only reason that any appreciable fringes were recorded was that the reference beam and Fourier transform were very precisely balanced. It would not have taken much of a miss-match to greatly reduce the diffraction efficiency of the recording.

The experiment was repeated with both the amplitude of the transform's d.c., and that of the reference beam, reduced to 7. This produced a maximum fringe intensity of 98 which was at the end of the linear region of the film curve. In this case, the saturation effects seemed to attenuate the d.c. term with respect to the higher frequencies. This produced a degree of edge enhancement in the reconstructed square pulse. The transform was then scaled to give a maximum intensity of 40.5; near the centre of the linear region. It was now found that the higher frequencies were being attenuated by the saturation at the low exposure end of the curve. However, the reconstructed result was much nearer to the original square pulse. Further reduction in intensity lead to a more severe attenuation of the higher frequencies. The original transform and the three modified versions are shown in figure 3.25, and the reconstructions of the four transforms in figure 3.26.

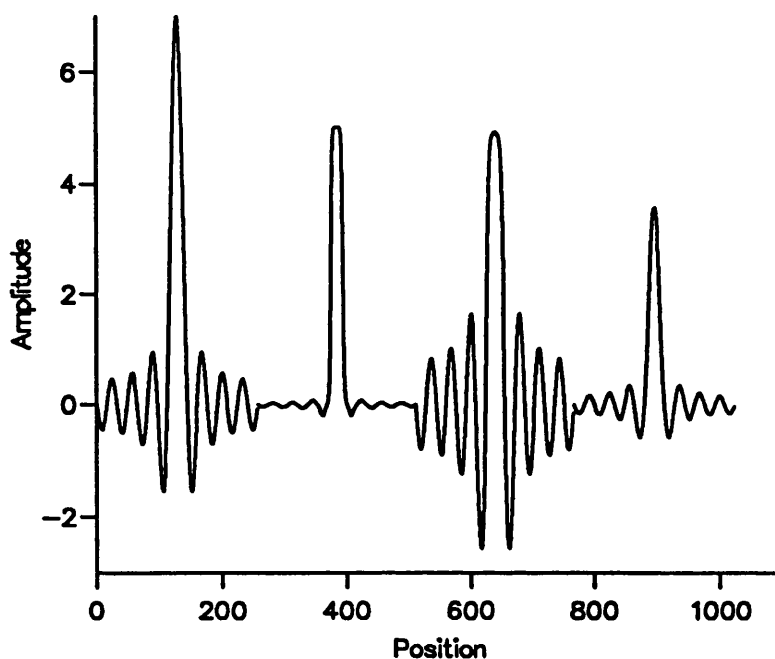


Figure 3.25: Fourier transform of a square pulse (left) and the modified transforms recorded with the reference beam balanced to the d.c. of the transform with the amplitude of the d.c. scaled to 16, 7 and 4.5.

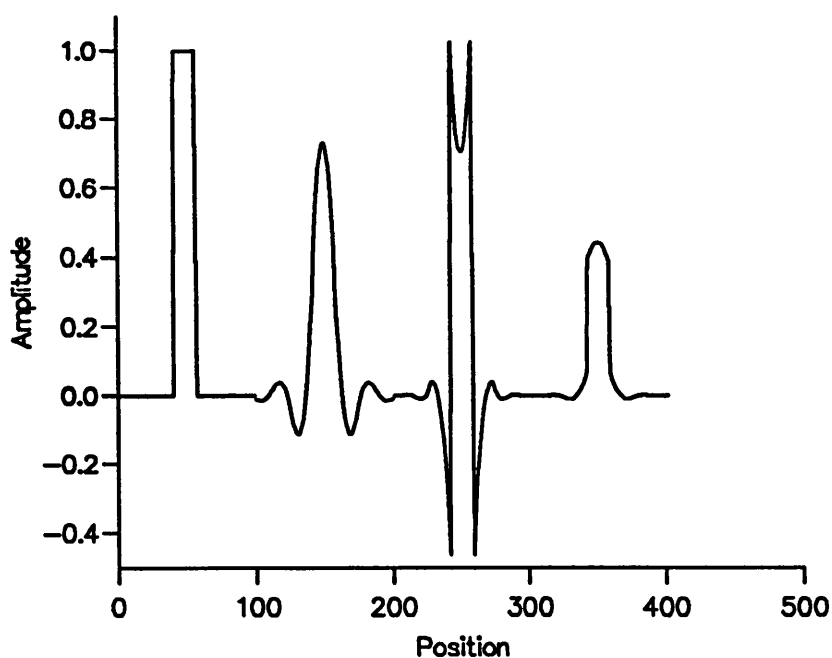


Figure 3.26: Reconstructions of the FTs from the previous figure.

A second set of experiments was performed with the reference beam balanced to the third lobe of the transform. The amplitude was scaled by factors of, 10.0, 2.5, 1.0 and 0.5. This moved the fringe intensity, at the point of highest modulation, from the high saturation end of the response curve down to the low saturation end. The modified transforms for the scaling factors 10 (maximum intensity in interference pattern 512), 2.5 (intensity 32), 1.0 (intensity 5.12) and 0.5 (intensity 1.28) are shown in figure 3.27. In this case, the use of high intensities lead to a very hard clipping of the d.c. values to a zero amplitude level. However, when very low intensities were used, the only significant recorded amplitude was at the d.c. Thus, a range of edge enhancements can be produced by a combination of absolute intensity and beam ratio manipulation. To confirm this point, the reconstructions of these transforms are shown in figure 3.28.

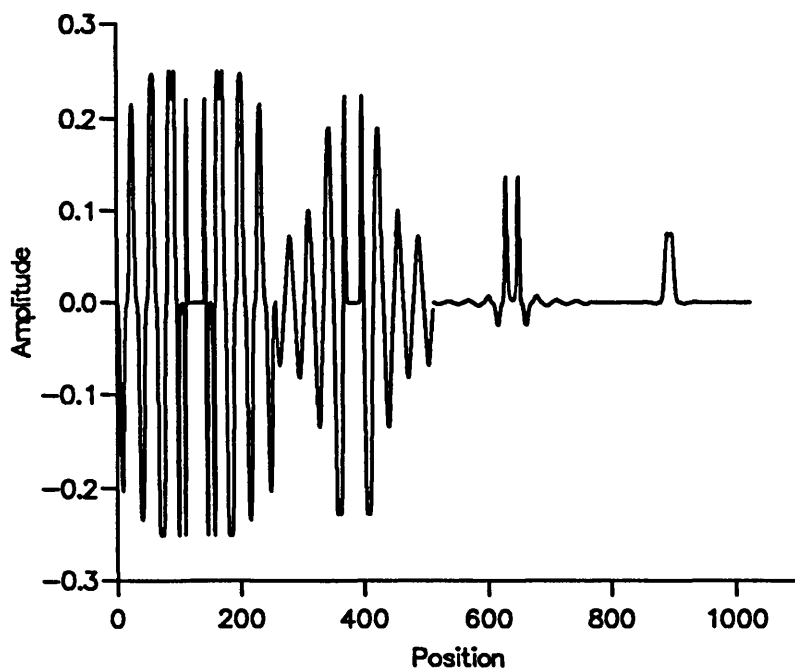


Figure 3.27: Modified transforms recorded with the reference beam balanced to the third lobe and with the original FT amplitude scaled by factors of 10, 2.5, 1.0 and 0.5

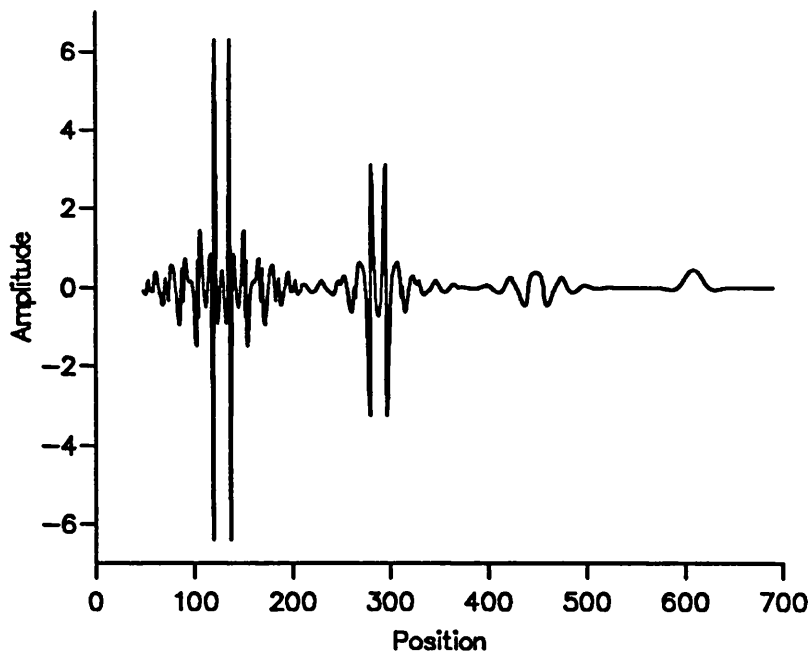


Figure 3.28: Reconstructions of the various FTs from the previous figure.

The full distorted fringe model produced results very similar to the amplitude model, apart from the absolute amplitude values. In order to apply the full fringe model, a 1-D FT of the fringe profile would have to be computed for every pixel in a Fourier transform. Bearing in mind this heavy computational load, it was decided that, in the light of the similar results produced, the amplitude model was sufficient. However, if it was necessary to take account of energy transfer through a system based on holographic film, the full distortion model could be used to predict the peak diffraction efficiency.

These experiments were repeated, using the BSO model to compute the stored FT amplitude. In this case, the absolute intensity did not affect the results. Figure 3.29 shows the modified transforms produced by balancing the reference beam to the d.c. and to the third lobe of the transform. The are shown in figure 3.30. It can be seen that, even with the reference beam balanced to the d.c. of the transform, there was still a slight degree of edge enhancement. It was found that increasing the reference beam further produced reconstructions that were closer approximations to the original square pulse. However, this was only achieved at the expense of reducing the diffraction efficiency of the hologram.

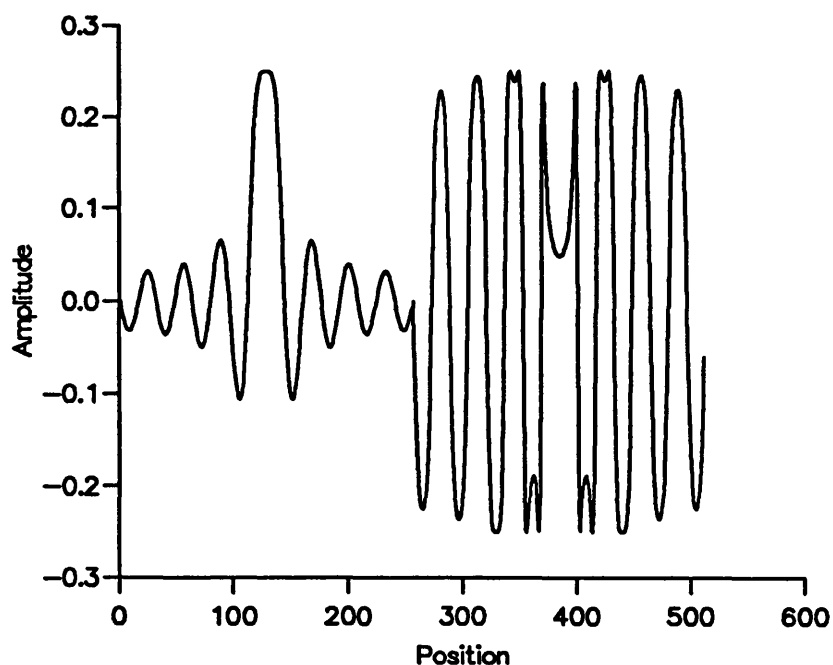


Figure 3.29: Modified transforms recorded with the reference beam balanced to the d.c. and the third lobe using the BSO model to compute the amplitude of the hologram.

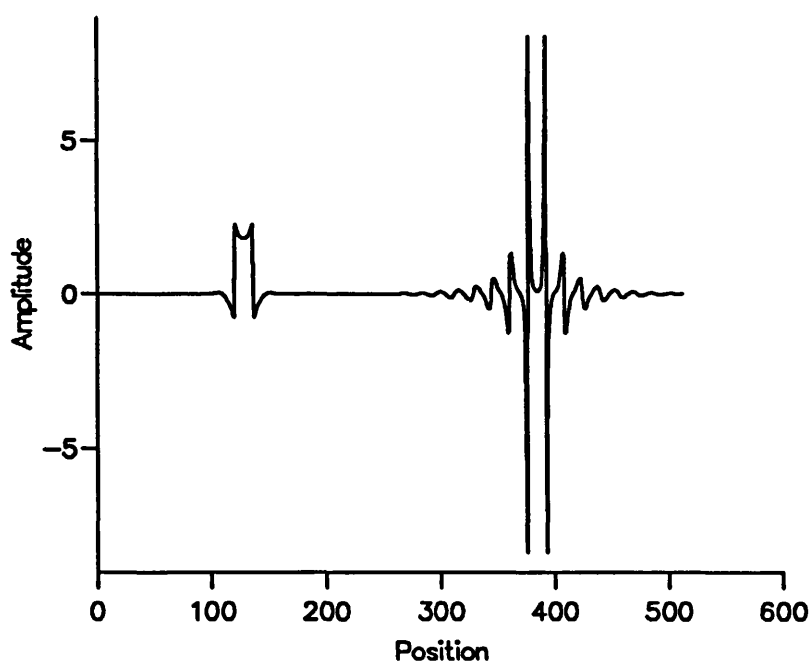


Figure 3.30: Reconstructions of the various FTs from the previous figure.

The main difference between the film and the BSO was the effect that the absolute intensities had on the film storage. The BSO could be used to record either 'perfect' or edge enhanced representations of the original image. However, unlike the holographic film, it was not possible to record in such a way as to attenuate the lower intensity portions of the FT with respect to the higher intensity areas. Such an attenuation will usually have a blurring effect as the lower intensities correspond to higher spatial frequencies. This property can affect the way that noise, in the reference image, is transmitted through a system. For example, figure 3.31 shows the test square with random amplitude noise added. Figure 3.32 illustrates the simulated results of recording and reconstructing the Fourier transform on film and BSO. The film recording was carried out at a relatively low absolute intensity in order to attenuate the lower intensity regions of the FT. This resulted in a reduced amount of noise in the recording compared to the BSO storage. However, as discussed in the introduction, in the case of BSO, an extra 'wash-out' beam can be used to spatially filter the recorded FT. Although often used for edge enhancement, this could also be used for any required blurring operations.

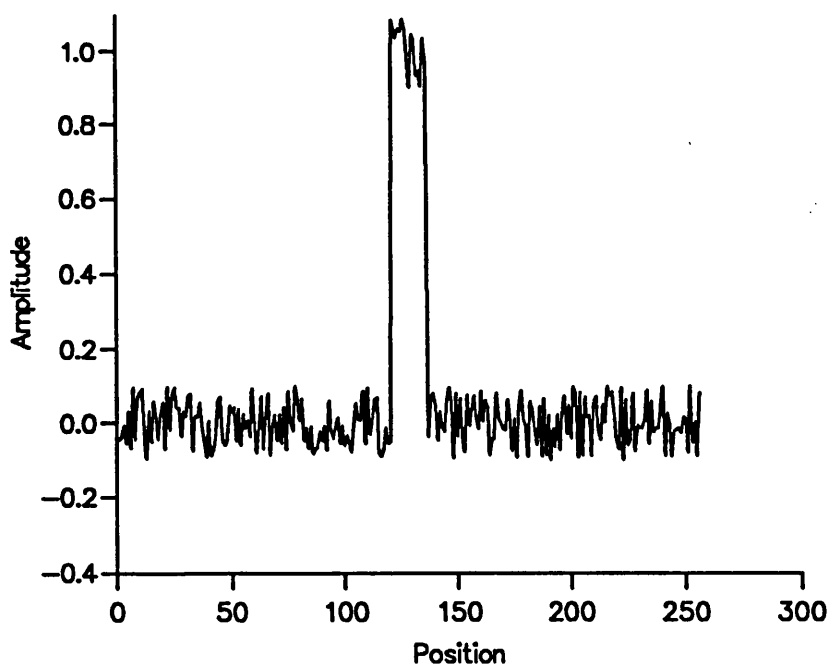


Figure 3.31: Square pulse test shape with random noise.

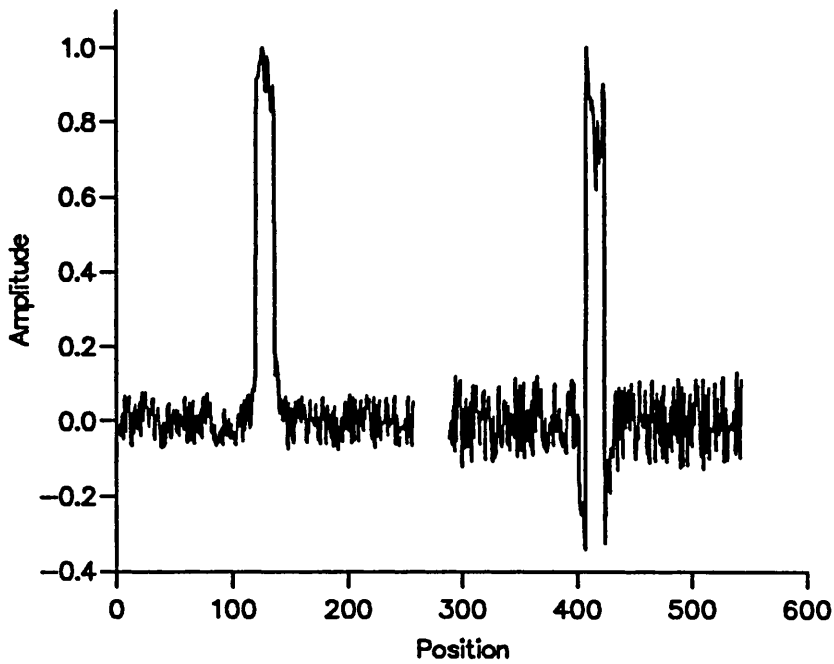


Figure 3.32: Reconstructions of the noise corrupted square pulse after simulated FT storage on holographic film (left) and BSO (right).

3.7 Beam Polarisation

Although not included in the work for this thesis, it is worth, for the sake of completeness, outlining the beam polarisation requirements for the correlator. For the proposed system, the grating vector and the applied field lie in the $\langle 110 \rangle$ crystal direction, i.e. in a direction normal to the 110 face. It is expected that the diffraction efficiency will be independent of the polarisation angle of a linearly-polarised, read-beam [1.74][3.12]. It is therefore advantageous to use an input polarisation angle that bisects the induced fast and slow axes of birefringence. In this case the light diffracted from the hologram is linearly polarised at 90° to the undiffracted beam. Some of the undiffracted light may be scattered into the correlation plane by defects in any of the system's optical surfaces. In this case, the signal to noise ratio can be significantly increased by using a polariser to block the scattered light and transmit the desired correlation signal [3.13]. The birefringence introduced by the applied field means that the light does not remain linearly polarised but is, in fact,

elliptical [3.14]. However, dramatic reductions in noise from scattered light can still be achieved in the correlator.

A brief experiment to measure diffraction efficiency as a function of read-beam polarisation was carried out. For larger fringe spacings, it was found to be reasonably constant. However, some variations were detected, which was consistent with experimental results reported in [1.74] and [3.12]. It was noted that the experiment was difficult to set up since many optical components, such as beam-splitters and wavelength filters, were found to impose their own polarisation selectivity on the results. An experiment to measure diffraction efficiency as a function of write-beam polarisation was also carried out and it was found to be fairly independent. In this case the optimum solution would be to use write-light of an orthogonal polarisation to that expected in the correlation plane, since a polariser transmitting the correlation plane would also help to block scattered write-light. However, a narrow-band, wavelength-selective filter would also provide an effective block to stop stray light from the writing beams, which were of a very different wavelength to the read-beams.

3.8 Conclusions

A reduction in the optical path lengths in the correlator would be beneficial on two counts. Firstly, the size of the overall system would be reduced, which is generally desirable when 'breadboarding' is considered. A second, and perhaps greater, advantage stems from reducing the path lengths on the write-side of the system. In this case, a shorter distance is travelled by the object and reference writing-beams before they are re-combined at the hologram. For the work being undertaken in this project, the only lasers available were continuous wave and not pulsed. This obviously puts a much greater emphasis on system stability. Even air currents can become a problem when the object and reference beam travel long distances to the hologram. Thus, reducing path lengths on the write-side of the system will relax the mechanical stability requirements.

In section 3.1, an expression for the location of the correlation plane has been derived as a function of the positions of the read and write image planes. If the

BSO hologram exhibits thin behaviour, the input planes can be brought close to their respective Fourier transforming lenses. Consider equation 3.12 for the case $\lambda_r f_r = \lambda_w f_w$: as d_r and d_w tend to zero; Z tends to f_c . Thus, an initial total path length of $2(f_r + f_w + f_c)$ can be reduced to $2f_c + f_r + f_w$. However, it was also shown that, if the hologram exhibits angular selectivity, complicated spatial filtering effects will be introduced if the images are not close to the front focal planes of the transforming lenses. This is because, if an image is away from the lens' front focal plane, the angles of the rays in the Fourier plane depend partly on the spatial frequency with which they are associated. The form of the spatial filtering is not a global function but changes with image position. The effects cannot, therefore, be alleviated with a transmission filter in the Fourier plane. This problem will be most pronounced on the read-side and it may be necessary to ensure that this image is close to the front focal plane of the lens.

In section 3.2, the lens and wavelength requirements were discussed. In order that it does not perturb the holograms, it is intended to use a read-beam of 633nm wavelength. The holograms will be written with light of 514nm, to which the BSO is relatively sensitive. In order to match the scales of the write and read Fourier transforms, the lenses must have a focal length ratio equal to the inverse of the write/read wavelength ratio. Furthermore, to minimise the spread of angles in the Fourier plane – while retaining all the image information – it is necessary to match the size of the transforms to the clear aperture of the BSO crystal. Achieving these criteria will, in general, require lenses of non-standard focal lengths. It was shown that such lenses could be constructed from a combination of two standard components. The focal length of the combination will be adjusted by altering the lens separation. It was also shown how an optimum lens combination can be selected, both with regard to the ease of setting up the system, and to the aberrations of the compound lens. To produce combinations with focal lengths in the 400 to 500mm range, it was found that a pair of 750mm achromatic doublets separated by a few hundred mm provided the optimum solution.

In section 3.3, methods were developed that allowed the computation of the inten-

sities in the correlation plane and in the Fourier plane. Furthermore, measurements of the BSO parameters, Θ^2 and τ_c , made possible the calculation of the response time; for any applied field, fringe spacing and input intensity. The powers of the lasers available for the construction of the correlator were known. Therefore the simulations, together with the known sensitivity of the CCD camera, would determine the required strength of applied field to produce a detectable correlation signal. Given the applied field, the fringe spacing, and the expected intensities in the Fourier plane; the time constant τ_g can be determined from equation 3.46.

A new consideration for BSO based systems, was discussed in chapter 3.4. If the holographic fringes become tilted with respect to the grating, this will cause a drop in diffraction efficiency. An expression was derived for the fringe tilts that will be produced in the correlator. Experimental results were also presented that measured the drop in diffraction efficiency with fringe tilt. One consequence of this effect was that, in order to avoid complicated spatial filtering effects, the write-images too, will have to be located at the front focal plane of their transform lens.

In the previous chapter, and in section 3.3, it was seen that it would be beneficial to use large fringe spacings. This was because of the decreases in angular selectivity and response time thus produced. In section 3.5 an expression for the maximum allowable fringe spacing was produced. Also, by considering the critical fringe spacing for the BSO, an upper limit on image size was found.

The 1-D simulation methods, developed in section 3.6, allowed a rapid, preliminary assessment of the effects of holographic storage on the amplitude of the Fourier transform. The storage in the BSO was modelled as a perfect phase recording. The amplitude was proportional to the fringe modulation between the FT and a holographic reference beam. The use of a reduced reference beam intensity was investigated, to provide varying degrees of edge enhancement. This produced results that were very similar to those produced by 'second order differentiation' filters such as the 'Difference of Gaussian'. These are characterised by the equal magnitude peaks, of opposite phase, that appear either side of the position of an edge in the original image. This type of edge enhancement is useful for correlation since it need

only be applied to one of the images being correlated. This is in contrast to first differential operators such as the Sobel filter, which need to be applied to both.

The main difference between holographic film and the BSO model was the effect that the absolute intensities had on the film storage. The BSO could be used to record either 'perfect' or edge enhanced representations of the original image. It was shown how this made noise suppression more difficult in the BSO case. However, an extra 'wash-out' beam can be used to spatially filter the FT recorded in the BSO, if required.

References: Chapter 3

- 3.1 Bage, M.J., Beddoes, M.P., "Lensless Matched Filter: Operating Principal, Sensitivity to Spectrum Shift and Third Order Holographic Aberrations", Appl.Opt., 15 p.2830 (1976).
- 3.2 "Optics", Hecht, E., Zajac, A., Addison-Wesely, 1974.
- 3.3 "Optics Guide 4", Melles Griot, commercial catalogue.
- 3.4 "Aberrations of Optical Systems", Welford, W.T., Adam Hilger, 1986.
- 3.5 Casasent, D., Luu, T., " Phase Error Model for Simple Fourier Transform Lenses", Appl.Opt., 17 p.1701 (1978).
- 3.6 Casasent, D., Luu, T., "Performance Measurement Techniques for Simple Fourier Transform Lenses", Appl.Opt., 17 p.2973 (1978).
- 3.7 Kidger Optics Software (IBM version).
- 3.8 Joyeux, D., Lowenthal, S., "Optical Fourier Transform: What is the Optimal Setup?", Appl.Opt., 21 p.4368 (1982).
- 3.9 "Principles of Optics", Born, M., Wolf, E., Pergamon Press, 1964.
- 3.10 Roberts, N.C., "Computer Simulation of the Effect of the Non-Linear Film Response on Optical Correlation", SRC tech memo. 203, (1985).
- 3.11 Thomas, C.E., "Film Characteristics Pertinent to Coherent Optical Data Processing Systems", Appl.Opt., 11 p.1756 (1972).
- 3.12 Marrakchi, A., Johnson, R.V., Tanguay, A.R., "Polarization Properties of Photorefractive Diffraction in Electrooptic and Optically Active Sillenite Crystals (Bragg Regime).", J.O.S.A.(B), 3 p.321 (1986).
- 3.13 Herriau, J.P., Huignard, J.P., Aubourg, P., "Some Polarization Properties of Volume Holograms in BSO Crystals and Applications", Appl.Opt., 12 p.1851 (1978).

- 3.14** Mallick, S., Rouède, D., Apostolidis, A.G., "Efficiency and Polarization Characteristics of Photorefractive Diffraction in BSO Crystals", J.O.S.A. (B), 4 p.1247 (1987).

Chapter 4

Digital Simulation of SLM Distortions

Contents of Chapter 4

- 4.1 Image Processing Methods 144
 - 4.1.1 Non-Linear Grey-Level Transfer 144
 - 4.1.2 Random Grey-Level Noise 149
 - 4.1.3 Structured Grey-Level Distortions 152
 - 4.1.4 Large Scale Phase Distortions 152
 - 4.1.5 Small Scale Phase Noise 161
 - 4.1.6 Signal Dependent Phase Distortion 164
 - 4.1.7 Device Resolution 167
- 4.2 A 2^N Factorial Experiment 168
 - 4.2.1 The Distortions Used in the Experiment 170
 - 4.2.2 Experimental Results 172
 - 4.2.3 Discussion of the Results 176
- 4.3 Conclusions 179

Abbreviations/symbols used in Chapter 4

BSO	Bismuth Silicon Oxide (Chapt. 1).
CCD	Charge Coupled Device (Chapt. 1).
CRT	Cathode Ray Tube (Chapt. 1).
d.c.	Zero-spatial-frequency component of FT (Chapt 1.).
LCLV	Liquid Crystal Light Valve (type of SLM).
MSF	Matched Spatial Filter (Chapt. 2).
MTF	Modulation Transfer Function.
SLM	Spatial Light Modulator (Chapt. 1).
D	Separation of two lenses in a compound lens (Chapt, 3).
f	Lens focal length.
F	Focal length of compound lens (Chapt. 3).
$N \times N$	Number of pixels in an image (Chapt. 1).
r_{im}	Half-width of input image aperture (Chapt.3).
λ	Wavelength.
Φ	Phase of a coherent light field.

After carrying out the preliminary work for the optical design of the correlator, problems were experienced with the acquisition of suitable spatial light modulators (SLMs). As a result of this, it was decided that the digital simulation work should be extended to model SLM performance. Many published papers on this subject were found; however, each of them dealt with one or two specific SLM characteristics. Furthermore, the applications and optimisations considered, were often either very specific or very vague. For example: [4.1] considers the effect of image grey-level noise on optically generated chord functions; while [4.2] considers the effects of SLM phase distortions on the width of the auto-correlation peak for the image of a girl's face. It was decided therefore, to concentrate on developing a set of computer programs to model various aspects of SLM performance. These programs would refer to a generic SLM rather than to a specific device. Once these programs were written, a suitable experimental method was chosen, to allow the effects of SLM distortions to be assessed for any given task. To prove the principle of the chosen method, the correlation task described in the introductory chapter was modelled with the inclusion of three image distortions.

4.1 Image Processing Methods

This section contains a description of the various algorithms that were employed in order to model various image distortions introduced by an SLM.

4.1.1 Non-Linear Grey-Level Transfer

This title covers any systematic discrepancy between the intensities in the image of a scene, as produced by a lens, and those relayed to the optical system. In order to produce a variable grey-level response, it was decided to construct a generic transfer curve with a set of distinct regions. This is shown in figure 4.1, together with the formulae used to create each region. The first of these was a uniform 'off-state' represented by $y = a$: where a is not necessarily equal to zero. A non-zero value may represent, for example, an imperfect linear polarisation produced by an SLM when the device is not switched on. This allows light to leak past the polariser that is

placed after the SLM; which is used to change the polarisation modulation, produced by the device, into an intensity modulation. Next, as the device begins to switch on, the slope of the curve increases at a rate governed by the constant b . The point at which the switching started was the specified input intensity, I_1 .

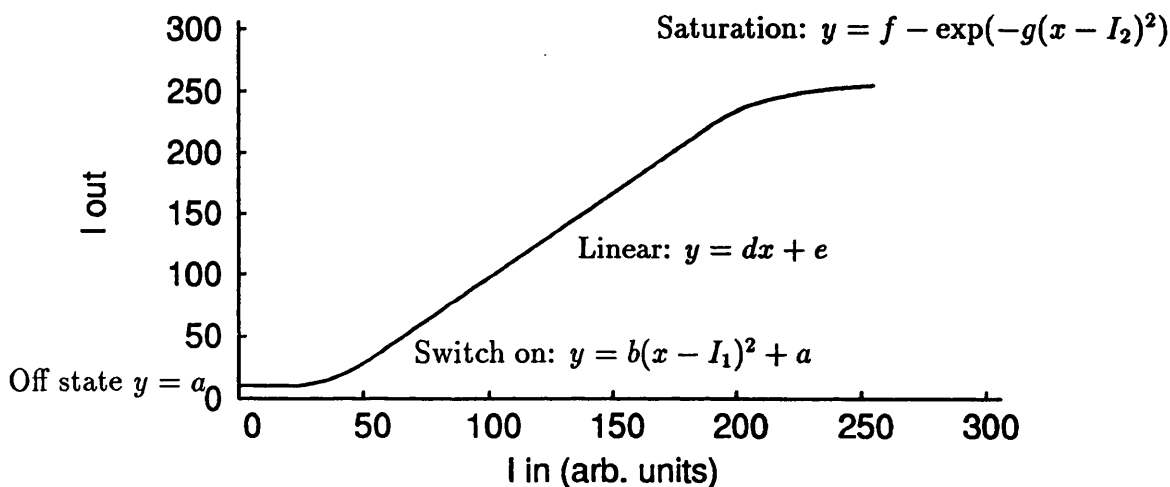


Figure 4.1: Generic grey-level transfer curve.

This switching process eventually leads to a linear region, of specified slope d . The computer program that produced the curve, automatically calculated the value of intercept for the linear region, e . This was chosen so as to join the linear part of the curve to the ‘switch-on’ part. Finally, at a specified intensity I_2 , the output begins to saturate. The rate at which saturation is reached, is determined by the specified constant g . The computer program set the value of the offset f to join the saturation curve to the linear region. Splitting the curve into separate regions in this way, made control of the various parameters easier.

To illustrate the effects produced by this sort of grey-level distortion, a cross correlation was performed between the images of two machined components: component#2 and component#3. These are shown in the top half of figure 4.2. Two sets of transfer curves were produced. The first, shown in figure 4.3, contained curves that converged to a thresholding operation at a level of 128. The second set converged to a threshold at a level of 40, as shown in figure 4.4. Two distorted versions of the image of component#2 are shown in the bottom half of figure 4.2. They were

produced using the most non-linear curve from each of the two sets. The main visual difference between the two original component images, was the structure present inside the boundary of the object. This was only visible in the image of component#2. This difference was enhanced by one of the sets of transfer curves, but was hidden by the second set.

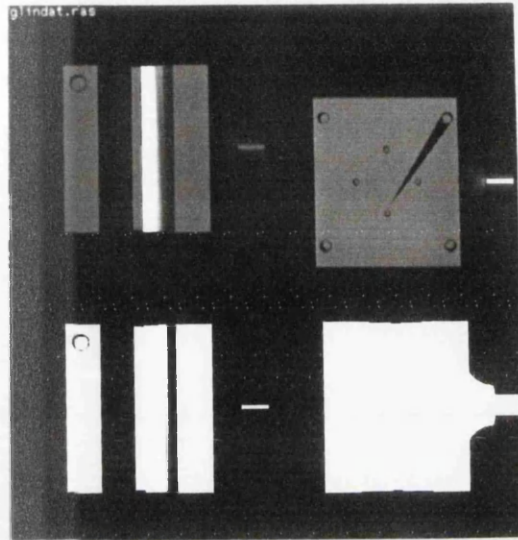


Figure 4.2: Perfect images of component#2 and component#3 (top) and two different grey-level distorted versions of component#2 (bottom).

A series of correlations was carried out using the two sets of transfer curves, which were applied to both of the images. The discrimination of the correlation process was defined as the ratio P_{auto}/P_{cross} : where P_{auto} is the magnitude of the peak of the auto-correlation function for component#2, and P_{cross} that of the cross-correlation peak formed by the two components. The images were edge enhanced prior to correlation. The variation in discrimination, as the different curves were applied, is shown in figure 4.5. Results for both sets of curves are plotted together. The triangles illustrate the changes as the transfer curves tend to a threshold at the level of 128; the circles plot the change from a near linear curve to a threshold of 40. In both cases, curve 0 is the result with no distortion applied and curve 5 represents the most non-linear transfer.

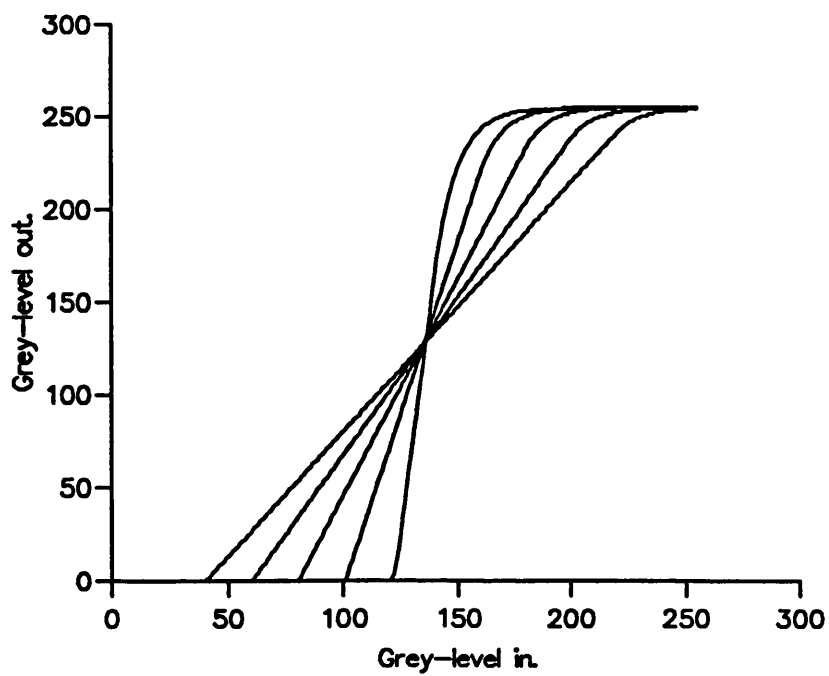


Figure 4.3: First set of transfer curves.

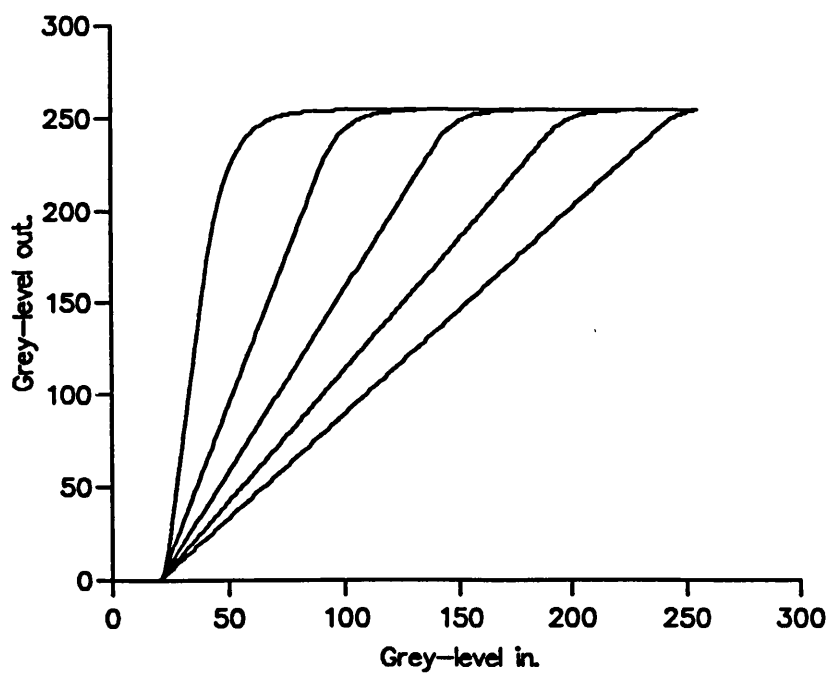


Figure 4.4: Second set of transfer curves.

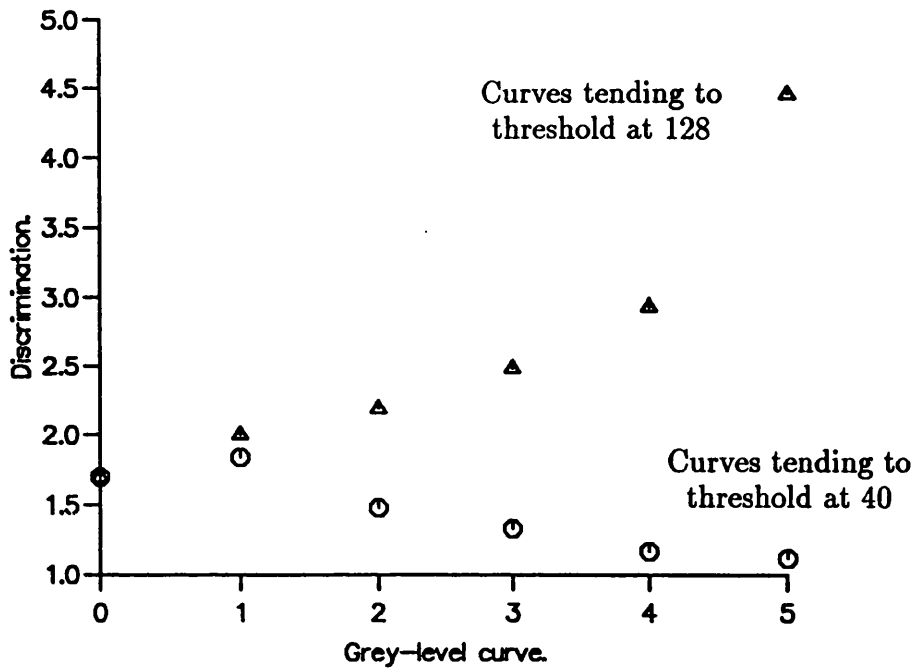


Figure 4.5: Correlation discrimination, between components#2 and #3, varying with the different grey-level transfer curves.

What these experiments showed, was that a very wide range of results could be produced, if the dynamic range of an SLM is small compared with that of the input imagery. Although the results, in some cases, appeared to be beneficial, this would be a false assumption. This is because, for all but the simplest tasks, a large amount of prior knowledge would have to be built into the system to allow it to set an appropriate threshold level. Thus, when assessing the effects of limited dynamic range, care must be taken to avoid drawing erroneous conclusions as a result of, for example, using human intelligence to set threshold levels. For example, research reported in [4.3] concluded that binary data could give improved signal to noise ratios in the correlation plane, compared to noisy grey-level data. However, no mention of the effort required to determine suitable thresholds was made. Furthermore, there was no discussion in the reference of the effects of the actual threshold level on the discrimination between different objects.

Given an SLM of limited dynamic range, the effects of grey-level transfer would need to be explicitly modelled for each task that was to be performed by the correlator. This would be especially important if was expected to distinguish between

objects that differed merely by virtue of details contained within the boundary of an object.

4.1.2 Random Grey-Level Noise

As well as the non-linear grey-level transfer, the intensities in an image displayed on an SLM may be corrupted by random intensity noise. This could be caused by:

1. Variations in the responses of the individual CCD elements in the camera viewing an object.
2. Local variations in the CRT phosphor brightness.
3. Variations in the local sensitivity of the photo-conductor on the SLM.
4. Imperfect alignment of the liquid crystal molecules in the SLM
5. Defects and blemishes on the optical components both before and after the SLM.

One form of noise, often used in optical processor simulations, is Gaussian additive noise [4.1] [4.4]. This noise can be shown to be associated with shot noise and thermal noise in a generic photo-detector [1.1]. In order to generate this Gaussian noise using a Fortran random number generator the following rule was applied:

$$I_{out} = I_{in} + A * \left(\frac{\sum_{a=1}^{a=n} \text{rand}(0)}{n} - 0.5 \right) \quad (4.1)$$

where A was the amplitude of the noise; and the Fortran function $\text{rand}(0)$ produced random numbers between 0 and 1. If I_{out} was greater than 255 or less than 0, it was clipped to these values. The width of the Gaussian function could be increased or decreased by averaging over smaller or larger n .

To show that this random number averaging produced a Gaussian distribution, an experiment was carried out in which 5000 amplitude values were generated. A was set to 1 and I_{in} was set to 0.5; so the values of equation 4.1 varied between 0 and 1. The resulting values were sampled with an interval of 0.01. The population of the intervals, for two different values of n , is shown in figure 4.6. The broad curve

was generated by using $n = 4$ in equation 4.1; the narrow one by using $n = 10$. Also shown are two ideal Gaussian curves and it can be seen that the averaged random numbers produced a good approximation.

To illustrate the effect of this noise, a corrupted version of the image of component #2 was produced. The parameters used were: $A = 510$ and $n = 10$. This meant that the maximum possible error at an individual pixel was ± 255 . To examine the effect of this noise on a correlation result, two, different, noisy versions of the input image were produced. The random number generator was started from a different seed point for each application of noise. This ensured that the noise in the two images was uncorrelated. One of the images was Fourier transformed and the amplitudes were modified to simulate storage in a BSO crystal; as described in section 3.6 of chapter 3. The reference beam amplitude was set to be 32 times less than that of the transform's d.c., in order to provide edge enhancement in the correlation process. One of the noise corrupted images is shown on the left of figure 4.7. On the right of the figure is the result of reconstructing the BSO-stored Fourier transform.

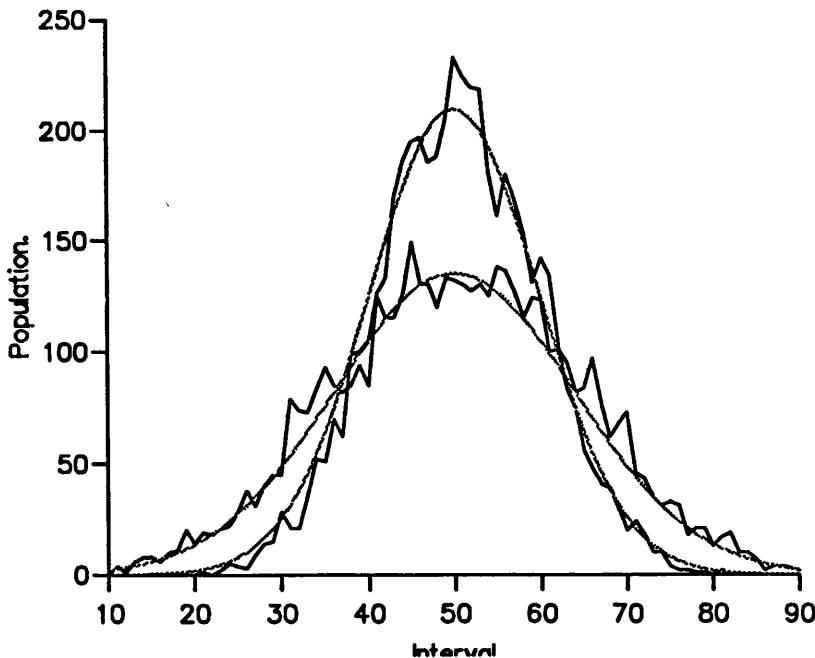


Figure 4.6: Gaussian populations, generated by averaging values from Fortran random number generator. The dashed lines represent perfect Gaussian curves.

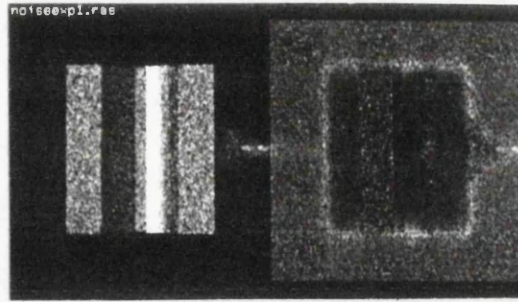


Figure 4.7: Noise corrupted image of component#2 (left) and the result of simulating its storage as a BSO hologram (right).

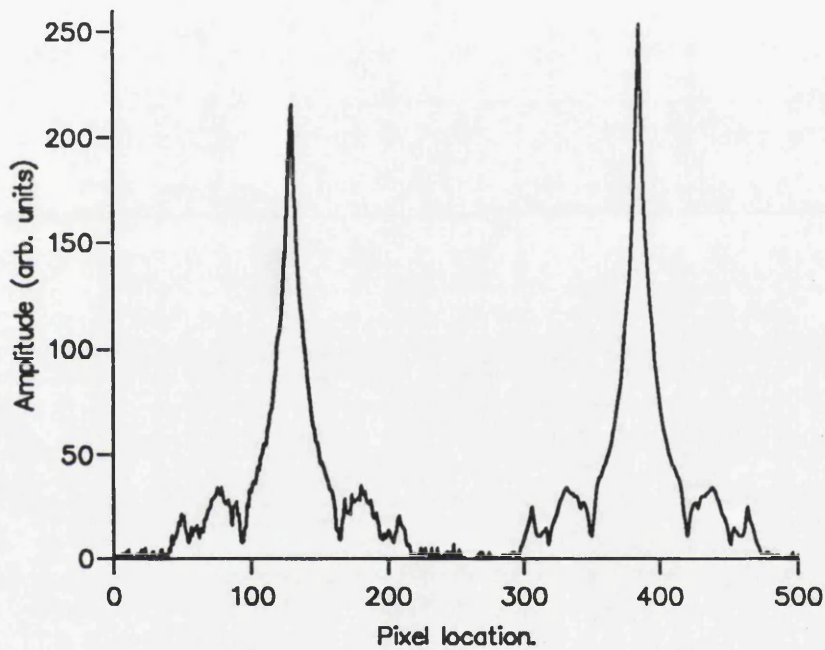


Figure 4.8: Corresponding slices through the noise-corrupted correlation plane (left) and the perfect auto-correlation (right).

Figure 4.8 shows a slice through the correlation peak, with a corresponding section through the un-corrupted auto-correlation for comparison.

From the experimental results, it can be seen that the random noise had very little effect on the correlation result. The profile of the corrupted peak is itself slightly noisy; but the original structure of the perfect auto-correlation is fully in tact. The reduction in the peak amplitude would correspond to a drop in intensity to about 70% of the auto-correlation peak value. However, it is unlikely that this would be

significant, especially considering the severity of the noise that was applied to the images. It is concluded that random grey-level noise of this type is unlikely make any serious impact on the correlation results.

4.1.3 Structured Grey-Level Distortions

For optically addressed SLMs, thickness variations across a device can cause both a phase distortion and a change in the grey-levels. Often, the most noticeable manifestation of this is non-uniformity in the dark off-state of the device. Developing a full model of this effect was not within the scope of this research. However, it was relatively straight forward to produce a program that would alter, in a realistic way, the lower grey-level values of an image.

The thickness of the SLM, T , may depart from perfect flatness - e.g. a uniform thickness Δ - according to:

$$T = \Delta + \delta(x, y) \quad (4.2)$$

Four different forms of $\delta(x, y)$ will be discussed in section 4.1.4; which deals with large scale phase distortions. The grey-level values of an image were then changed from $I_{in}(x, y)$ to $I_{out}(x, y)$ according to:

$$I_{out}(x, y) = \begin{cases} I_{in}(x, y) + (A/2) + (A/2) \cos(\delta(x, y)/\alpha) & \text{if } I_{in} < A \\ I_{in}(x, y) & \text{if } I_{in} > A \end{cases} \quad (4.3)$$

The level A determined the severity of the effect by setting the level below which the pixel values would be altered. The constant α , merely determined the number of cycles across an image, for a given variation in device thickness.

This form of distortion was one of those included in the 2^N factorial experiments, reported in section 4.2.

4.1.4 Large Scale Phase Distortions

These distortions are caused by thickness variations across the SLM; and are independent of the input signal. Reported phase distortions can range from less than $\lambda/4$ - for high quality research SLMs [4.5] - to many wavelengths for cheap, commercially available devices [4.6]. It should be pointed out that the effects of the

phase distortion can be alleviated, or cured completely, by using additional optical elements in the system. These elements include phase conjugating holograms [4.6] and liquid gates [4.7]. It has also been proposed, that programmable, SLM-based, Fourier plane filters could be used. These would adapt themselves to produce optimum signals in the presence of phase distortions [4.8]. Another method suggested to overcome the problem was the encoding of the conjugate of the phase distortion onto the SLM, using the phase modulating properties of the liquid crystal layer [4.9]. However, these measures all imply increased system complexity and, in the case of holograms, reduced light efficiency and increased noise in the optical system. Therefore, they may be regarded as undesirable; unless the cost of manufacturing devices to the required specification is prohibitively expensive.

In order to model the effects, a computer program was written to add phase distortions to input images, which could then be correlated. Studies such as these have been carried out before [4.2], but as pointed out in this reference, the results are always application and image specific. At the input plane, the images were initially considered to have zero phase. The first distortions that were modelled were the most obvious cases: a ramp distortion, and a spherical distortion. The ramp distortion involved altering the phase of the image from $\Phi_{in}(x, y)$ to $\Phi_{out}(x, y)$ according to:

$$\Phi_{out}(x, y) = \Phi_{in}(x, y) + \frac{x}{N} \times A_r \pi \quad (4.4)$$

where the image is assumed to be $N \times N$ pixels. The severity of the distortion was controlled by the constant, A_r . A computer generated interferogram of this function is shown in the top right of figure 4.9; together with an inset intensity representation of the phase.

The spherical distortion was created by applying a function which represented a sphere of radius $A_s \pi$:

$$\Phi_{out}(x, y) = \Phi_{in}(x, y) + A_s \pi - \sqrt{A_s^2 \pi^2 - x^2 - y^2} \quad (4.5)$$

The interferogram of this function, together with an inset intensity representation, is shown at the top left of figure 4.9. As expected, the addition of these phase terms to

an image, resulted in a degradation of the correlation peak; when a cross-correlation was performed with a perfect image. However, at this point, it is important to assess the significance of these results for an actual optical system.

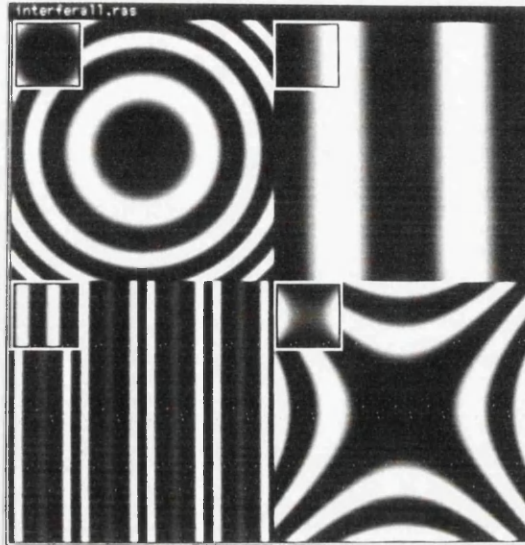


Figure 4.9: Computed interferograms of four phase distortions with different (x, y) variations. An intensity representation of the phase is inset in the top left corners of each of the interferograms.

If a phase ramp is added to an input image, this is equivalent to placing a glass wedge just after the SLM. This has the effect of deviating the beam away from the optical axis. This in turn leads to the light being focussed at an off-axis point in the Fourier plane; as illustrated in figure 4.10. It was found that the digital correlation could be recovered by translating the reference Fourier transform by an appropriate amount. This was not discussed in [4.2] or [4.9], although the phase distortions considered, were dominated by a ramp function. The distances involved may be very small. If the phase ramp causes a change of $A_r\pi$ across the width of the input image, the angle of deviation caused will be:

$$\alpha = \frac{A_r\pi}{2r_{im}} = \frac{A_r\lambda}{4r_{im}} \quad (4.6)$$

where r_{im} is the radius of the image. The displacement in the Fourier plane is then found, from equation 2.29, to be: $\delta = f\alpha$. For a 400mm lens, with: $\lambda = 633\text{nm}$, $A_r = 4$, and $r_{im} = 5.0$; the displacement is found to be: $\delta = 50\mu\text{m}$. Nevertheless,

this can lead to a dramatic reduction in the magnitude of the correlation peak. Figure 4.11 is a plot of the correlation plane for an auto-correlation of component#2. The correlation was carried out using a simulation of a BSO correlator; the ratio of the amplitudes of the reference beam and the d.c. being 1/16. The effect of the 4π phase ramp, on the correlation peak, is illustrated in figure 4.12. It is seen that the correlation has failed to locate the object.

Image + Phase ramp

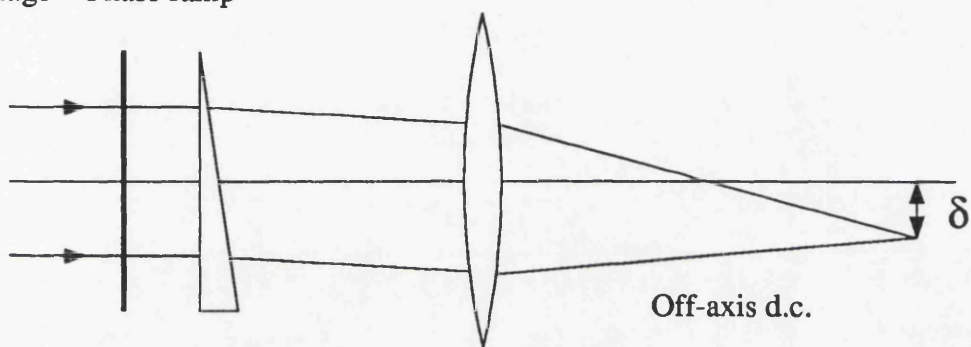


Figure 4.10: Effect of a phase ramp on the position of the Fourier plane.

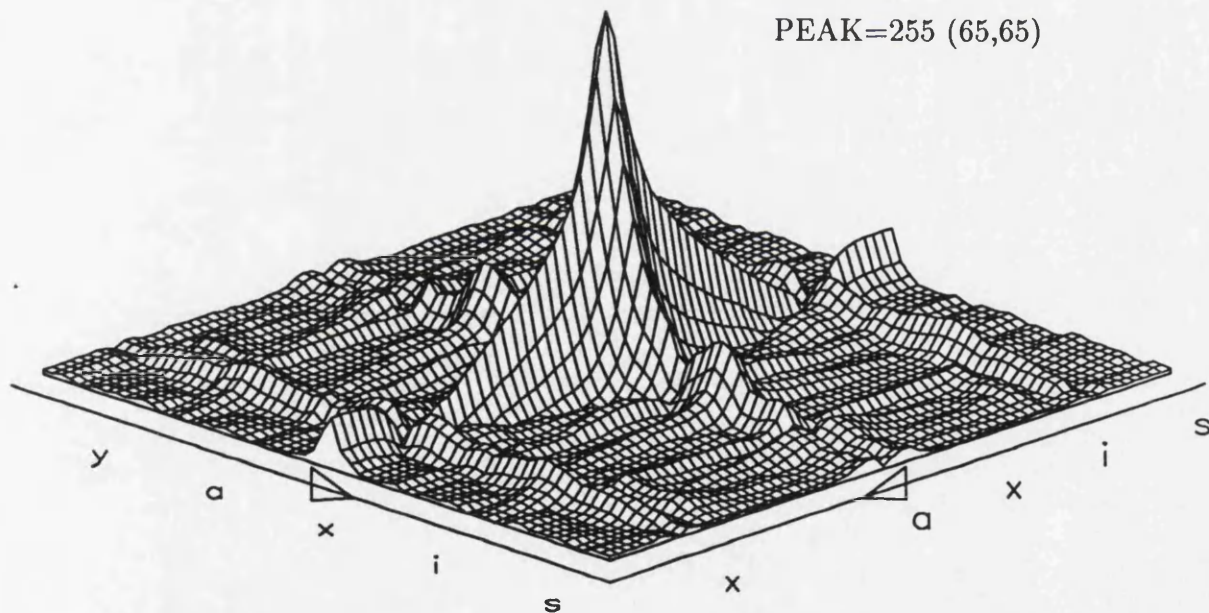


Figure 4.11: Auto correlation of component#2: BSO correlator with 1/16 beam amplitude ratio

PEAK=119 (65,104)

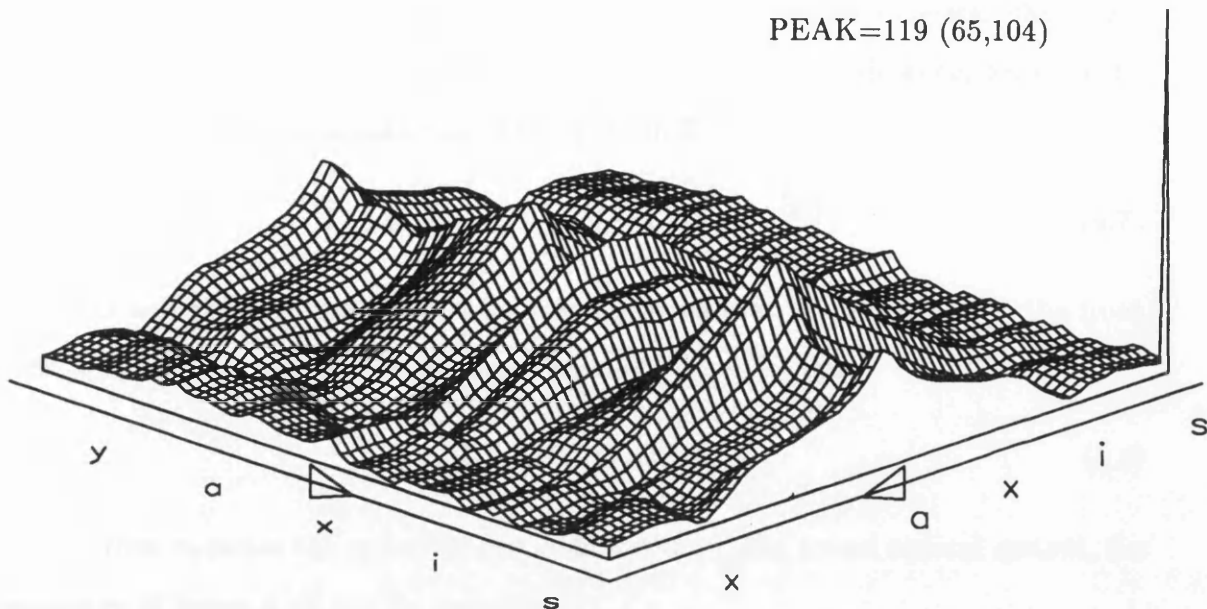


Figure 4.12: Cross correlation of undistorted image of component#2, with second image with 4π phase ramp added.

Using a similar argument, the effect of adding a spherical distortion to the input plane, is equivalent to placing an additional lens just after the SLM. This has the effect of moving the location of the Fourier plane along the optical axis; as illustrated in figure 4.13.

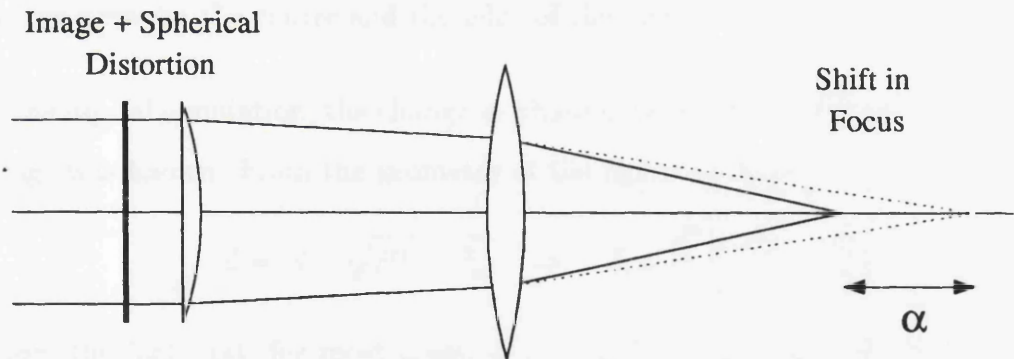


Figure 4.13: Equivalent effect of spherical phase distortion at the SLM.

Equation 3.16, in chapter 3, gave the focal length, F , of a system of two lenses of focal lengths f_1 and f_2 , separated by a distance D . In the correlator, the SLM will often be in the front focal plane of the transforming lens; thus, $D = f_2$. However, in this case, the equation predicts that the focal length of the system is equal to

the focal length of the second lens, f_2 . This implies that the spherical phase will not alter the apparent focal length of the transform lens. The distance, between the second lens and the focal plane is; from equation 3.17:

$$F + O_2H_2 = F - \frac{FD}{f_1} = f_2 \left(1 - \frac{f_2}{f_1}\right) \quad (4.7)$$

where D and F have been substituted for; assuming that the image was at the front focal plane of the lens f_2 . Thus, α , in figure 4.13, is given by:

$$\alpha = f_2 - (F + O_2H_2) = \frac{f_2^2}{f_1} \quad (4.8)$$

In order to relate the spherical phase distortion to the actual optical system, the geometry of figure 4.14 can be considered.

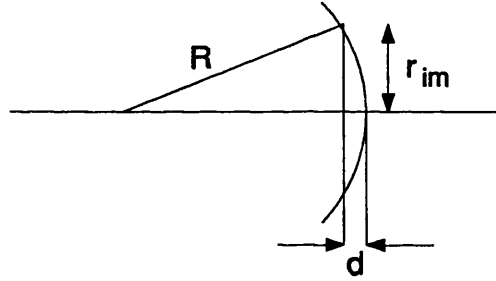


Figure 4.14: Geometry for calculating the apparent focal length from the phase change between the centre and the edge of the field.

In the digital simulation, the change in phase between the centre and the edge of the image was known. From the geometry of the figure we have:

$$d = R - \sqrt{R^2 - r_{im}^2} \Rightarrow R = \frac{d^2 + r_{im}^2}{2d} \approx \frac{r_{im}^2}{2d} \quad (4.9)$$

where the fact that, for most cases, $d \ll r_{im}$ has been used. If the phase change between the centre and the edge of the image is $\Delta\Phi = n\pi$ then:

$$d = \frac{\Delta\Phi}{2\pi} \times \lambda = \frac{n\lambda}{2} \quad (4.10)$$

This leads to an expression for R – which is equivalent to the focal length f_1 – of:

$$R = f_1 = \frac{r_{im}^2}{n\lambda} \quad (4.11)$$

For a value of $n = 1$; with $f_2 = 400\text{mm}$, $\lambda = 633\text{nm}$, and $r_{im} = 5\text{mm}$: it is found that $\alpha = 4\text{mm}$. The correlation plane, resulting from adding a spherical phase term to an image of component#2 is shown in figure 4.15. The peak is reduced in height, and the amplitude of the background structure in the correlation plane increased.

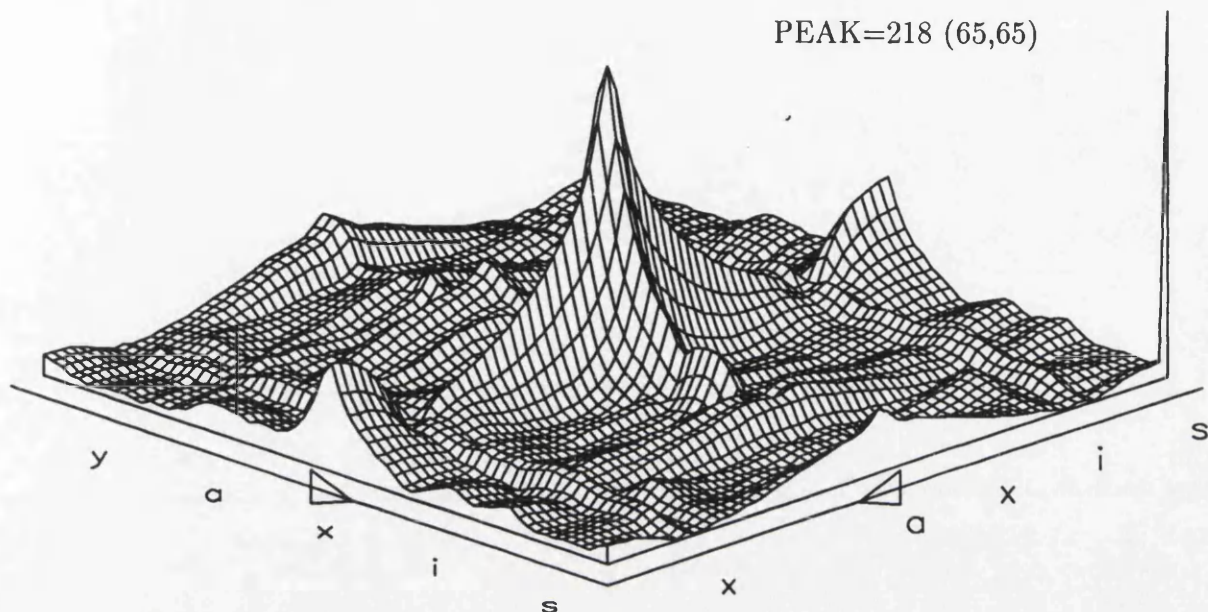


Figure 4.15: Cross-correlation of undistorted image of component#2, with second image with π spherical phase front added.

The application of either spherical or ramp phase distortions to an image, is equivalent to modelling the effects of misalignments in the Fourier plane. If an SLM suffers either a perfect ramp or a spherical phase distortion, the optical correlator may not be greatly affected. This is because any Fourier plane filter would be placed at the measured position of the lens focus, and not its theoretical location. Furthermore, if a measured distortion is found to have a ramp-like or a spherical component, the results of computer simulation will be over pessimistic. They could be improved in the equivalent optical system by repositioning the Fourier plane filter.

To simulate the effects of measured SLM phase distortions on the performance of an optical system, the following procedure should be adopted:

1. Measure the actual phase distortion of the device.
2. Subtract or add spherical phase profiles to minimise overall distortion.

3. Subtract or add ramp phase profiles to minimise overall distortion.
4. Perform simulations using the modified phase distortion.

Steps 2 and 3 are equivalent to refocussing the optical system and repositioning the Fourier plane filters respectively.

In order to continue the work on large scale phase distortions, new profiles were generated. These contained no overall spherical or ramp-like components. Two such functions are an 'astigmatic' function formed from two back-to-back half-cylinders and a sinusoidally varying distortion.

The cylindrical distortion was generated using:

$$\Phi_{out}(x, y) = \Phi_{in}(x, y) + A_{c1}\pi - \sqrt{A_{c1}^2\pi^2 - (N/2 - x)^2} - A_{c2}\pi + \sqrt{A_{c2}^2\pi^2 - (N/2 - y)^2} \quad (4.12)$$

The sinusoidal function was generated by:

$$\Phi_{im} = \Phi_1 + A_{sin} \sin\left(\frac{2\pi nx}{N}\right) \quad (4.13)$$

where n is the number of phase cycles across the image.

The interferogram of the sinusoidal and cylindrical functions, together with their intensity representations can be seen in the lower half of figure 4.9 on the left and right respectively. These profiles were generated using a sinusoid with an amplitude of 2.5π and two cylinders with radii 250π .

So far, the work had concentrated to the case where distortions were applied to one image only. This was reasonable since in the BSO correlator, different SLMs – with different distortions – would be used on the write and read-sides. In a conventional Vander Lugt correlator, employing only one SLM, the main effect of phase distortions is to destroy the shift invariance of the system. An MSF may be recorded with, say, the object at the centre of the image displayed on an SLM. If there is a phase profile associated with the SLM, then the region of this profile which overlaps the object will also be recorded. If an auto-correlation is formed, with the object still at the centre of the image, the phase conjugation properties of the MSF will eliminate the phase effects. However, if the object moves to a different part of the

input field, at another position on the SLM, it will see a different phase profile. The difference between this new profile and that recorded on the MSF will degrade the correlation peak.

An example of this loss of shift invariance was provided by correlating the two images shown in figure 4.16. The perfect correlation plane is shown in figure 4.17. It was formed by simulating a BSO correlator with a reference beam to d.c. amplitude-ratio of 1/16. An identical phase distortion, of cylindrical form, was then applied to both images. The correlation plane that was produced, is shown in figure 4.18.

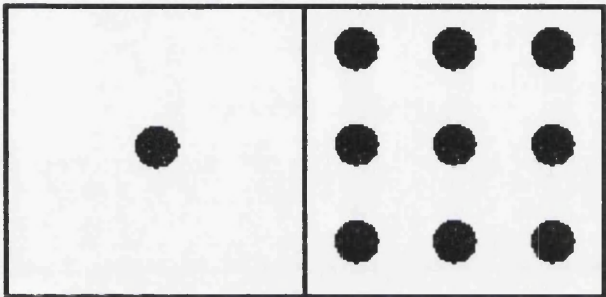


Figure 4.16: Input images used in the position invariance experiment.

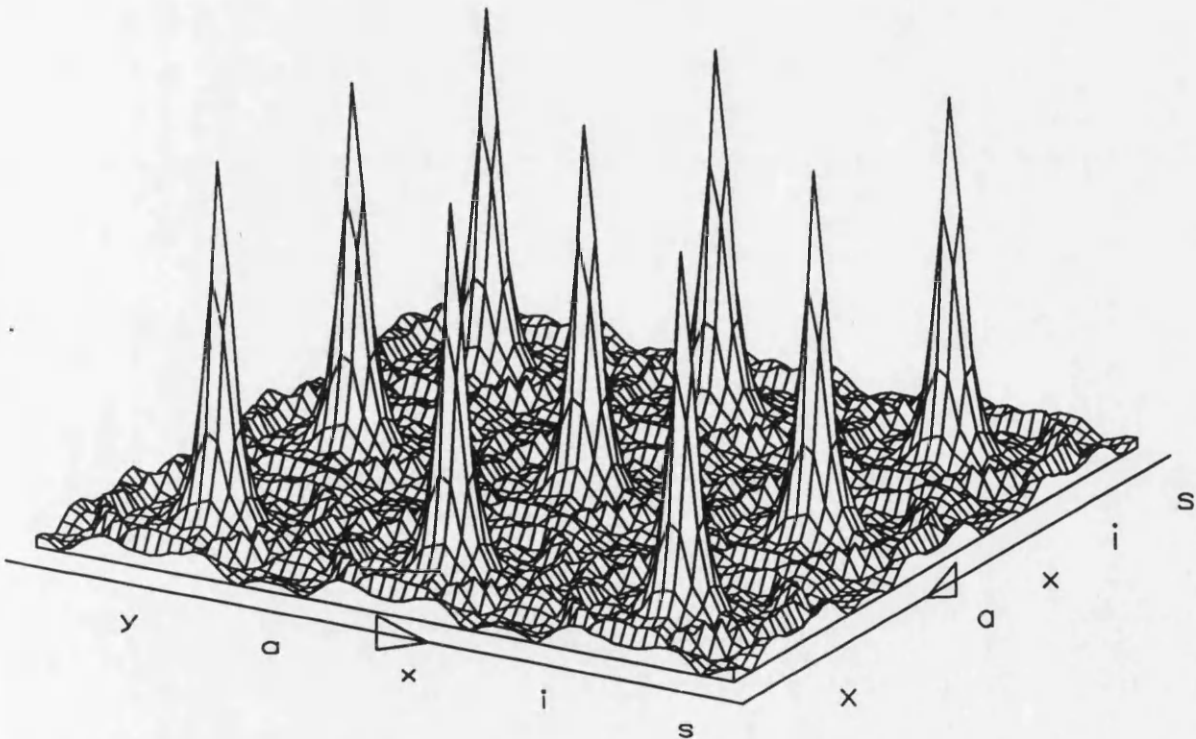


Figure 4.17: Correlation of single, centred hole template with an array of identical objects.

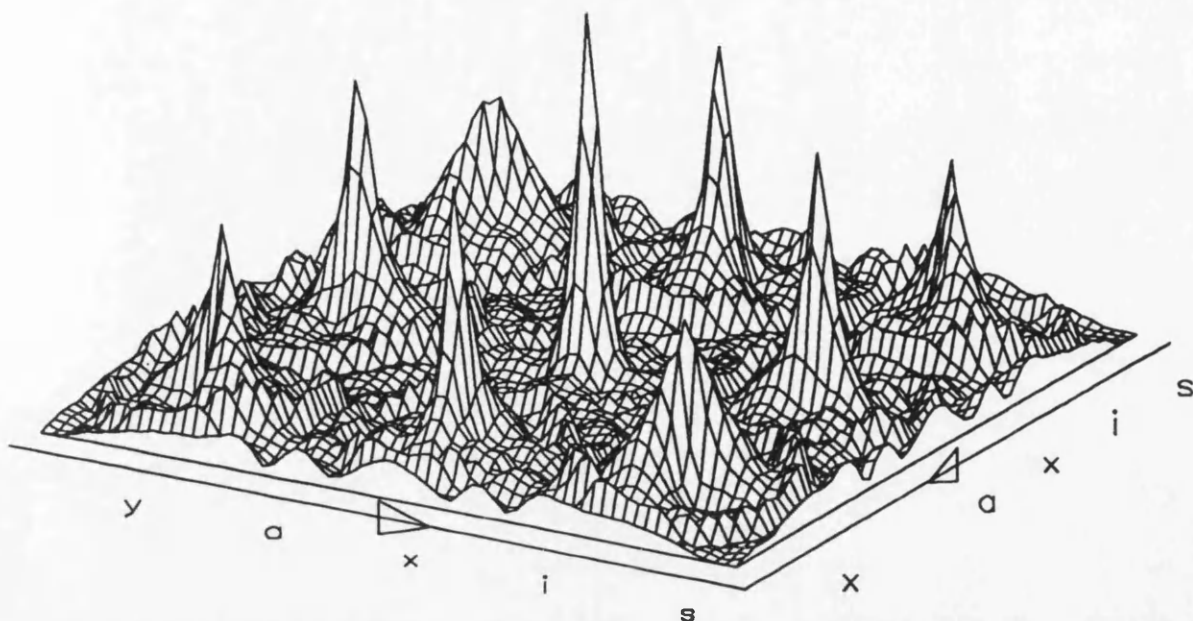


Figure 4.18: Correlation of hole template with array of holes, with identical, cylindrical phase distortion added to both images.

As expected those circles not at the centre of the image have suffered a drop in their correlation signals; i.e. the system has lost its position invariance.

Large scale phase distortion is one of the effects included in the experiment reported later in this chapter.

4.1.5 Small Scale Phase Noise

Previous work [4.10] had highlighted the problems caused by small scale phase distortions on the outputs of SLMs such as the Hughes LCLV. These distortions are caused by point defects in the layers that underlie the device's back mirror [4.11]. This, coupled with their random nature, meant that they can only be successfully eliminated at the device fabrication stage. Theoretically, these phase effects could be compensated for with a phase conjugating hologram. However, the phase pattern is at such a small scale that it would be impossible to align the compensating element with the SLM. These phase defects have been found to introduce a large amount of noise into optically edge enhanced images [4.10]. They will also produce noise in the Fourier plane, by random scattering of incident light. Papers such as [4.12], in which a Hughes SLM was used, make no mention of the problem although it should have

had an effect on the results. In the absence of any reported data, a brief experiment was performed in order to provide some numbers for the simulation work.

A Twyman-Green interferometer was set up to provide a magnified image of the back mirror of a Hughes LCLV. In this system, the image was interfered with the beam from a second mirror. Two images were then acquired: one with two mirrors in the system; and one with the SLM and one mirror. The images were stored digitally which allowed the fringe noise to be highlighted by contouring the fringes. The results are shown in figure 4.19; the mirror fringes are on the left and the SLM fringes on the right. The field of view is 2.15mm and the defect size was found to be around 0.1mm. This equated to about 4 lines on the CRT that would have been used in the correlator. The peak distortion appeared to be about $\lambda/8$.

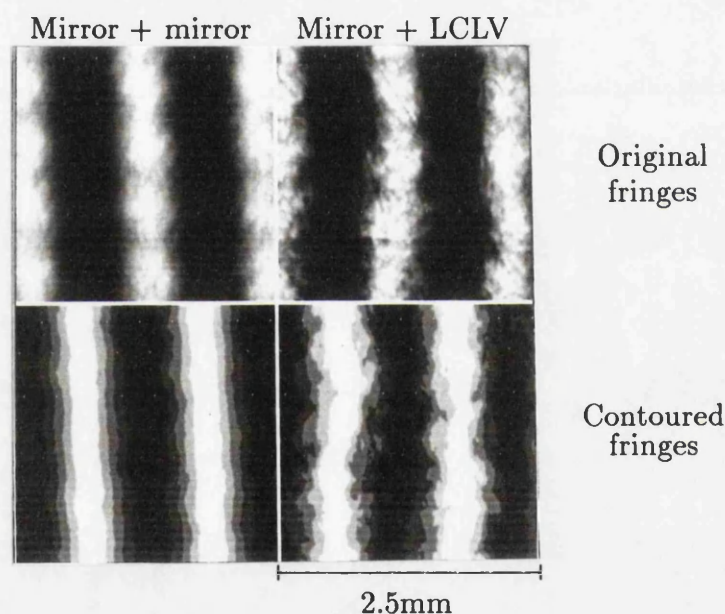


Figure 4.19: Comparison of mirror and LCLV fringes

In order to generate noise of this type, a program was written which randomly scattered phase noise 'templates' over an image. These templates were 1 pixel, 3x3, and 5x5 pixels in size. The magnitude of the phase of each template was also randomly varied between specified maximum and minimum values. When the program was run, the number of defects of each size was specified by the user. For each template, of each size, up to the maximum number specified, the program generated random x, y locations and a random amplitude. The phase template was

then added to the phase already existing at the x,y location. For the 3x3 and 5x5 pixel templates, the phase values added to the central pixel and its neighbours were:

$A\pi \times$	0.26	0.51	0.26
	0.51	1.0	0.51
	0.26	0.51	0.26

$A\pi \times$	0.15	0.30	0.38	0.30	0.15
	0.30	0.61	0.79	0.61	0.30
	0.38	0.79	1.00	0.79	0.38
	0.30	0.61	0.79	0.61	0.30
	0.15	0.30	0.38	0.30	0.15

where $A\pi$ was the amplitude of the noise at the particular location. These templates approximated Gaussian profiles. They were designed to ensure that, in adding noise templates to the image, no sharp edges were introduced.

An experiment was carried out to assess the effects of this phase noise on an edge enhancement process. A 256×256 pixel image of component#2 was corrupted with phase noise. The storage of its Fourier transform, as a BSO hologram, was then simulated. The holographic reference beam was given an amplitude that was sixteen times lower than that of the transform's d.c.; this produced an edge enhanced image on reconstruction. The noise that was added to the image consisted of 5000 single-pixel defects, 10000 3x3 pixel defects and 5000 5x5 pixel defects. The maximum amplitude of any defect was 0.25π ; i.e. $\lambda/8$. The results are shown in figure 4.20.

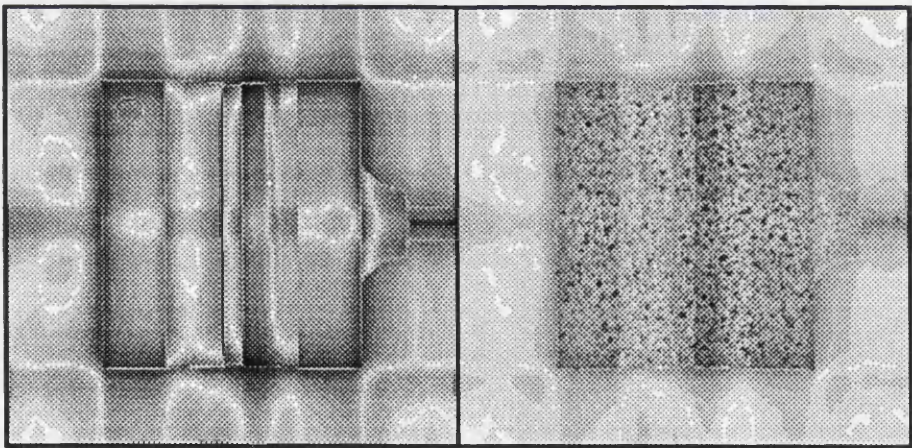


Figure 4.20: Simulated results of edge enhancement, by BSO holographic storage, for images with (right) and without (left) phase noise.

The edge enhancement of an undistorted is shown on the left which may be compared with the enhanced, noisy image on the right. This demonstrated how the edge

enhancement process highlights the phase noise structure, at the expense of the edges in the image.

This type of SLM distortion was further investigated in the experiment reported later in this chapter.

4.1.6 Signal Dependent Phase Distortion

Several papers have been published on the cross-coupling between amplitude and phase in the output from liquid crystal based SLMs. Reference [4.13] considered the case of a twisted nematic cell between crossed polarisers. It suggested that the main phase modulation took place at lower voltages than those required to affect the amplitude of the output. No observable cross-coupling would occur, since the phase output was reportedly constant at the voltages used for amplitude modulation. In [4.14] however, it was stated that a linear amplitude-phase relationship, with some quadratic behaviour, existed in the output of a Hughes LCLV. In the simulation work performed here, the requirement was to study generic systems and not specific devices. It was therefore decided, to model the cross-coupling as a linear relationship, although this could easily be extended to include higher order terms.

The input image might already contain phase information, for example if the image had already been subjected to any of the phase distortions discussed above. If this was not the case, then the image array was assumed to have zero phase initially. In either case, the phase of the output image at each pixel was given by:

$$\Phi_{out}(x, y) = \Phi_{in}(x, y) + A\pi\sqrt{I_{in}(x, y)/I_{max}} \quad (4.14)$$

where the square root of the input intensity was taken to give the amplitude values. $A\pi$ was the maximum phase distortion that the device was capable of producing and I_{max} was the predetermined, maximum image grey-level. In order to illustrate this effect, for an SLM in the input plane of an optical system, the edge enhancement of the image of component#2 was considered. In this case, the enhancement was performed by a Fourier plane band-pass filter. A maximum, signal dependent phase distortion of 2π was used; i.e. this would be the phase added to an input pixel of grey-

level 255. The results of the edge enhancement are shown in figure 4.21. The image has been negated, and the square root of the output intensity taken, to compress the dynamic range for printing.

It was clearly seen that the phase distortion greatly enhanced the internal detail of the test object relative to its external outline. It also increased the magnitude of the computed auto-correlation function by a factor of 1.5 times. Predictably, the actual shape of the correlation peak was altered as the peak became elongated. This was due to the stronger vertical lines in the edge enhanced image.

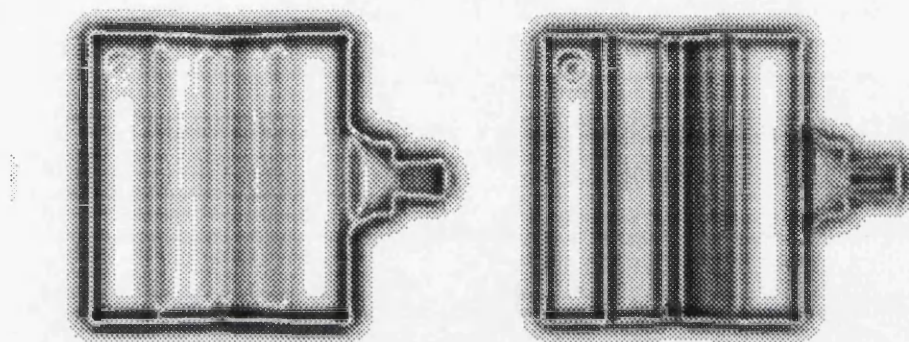


Figure 4.21: Effect of amplitude-phase cross coupling on edge enhancement.

A second experiment was carried out, to illustrate the effect of signal dependent phase distortion on the discrimination of a correlation system. The two images shown in figure 4.22 were cross-correlated. A reference template was made of component#2, which was then correlated with a second image, containing component#2 and two others. The correlation was first performed on the images, assuming that they both had no phase associated with them. The resulting correlation plane is shown in figure 4.23.

Both images were then subject to a phase distortion, with a maximum of 2π , and the correlation repeated. The result of this is shown in figure 4.24. Both a large increase in discrimination and a change in correlation peak shape can be clearly seen. The results of this simulation were very interesting. The presence of amplitude-phase cross-coupling was shown to increase the discrimination between objects in a correlation experiment.

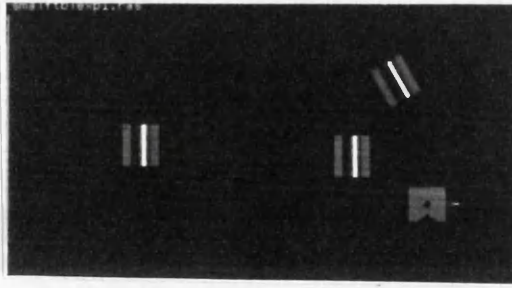


Figure 4.22: Input data for discrimination experiment.

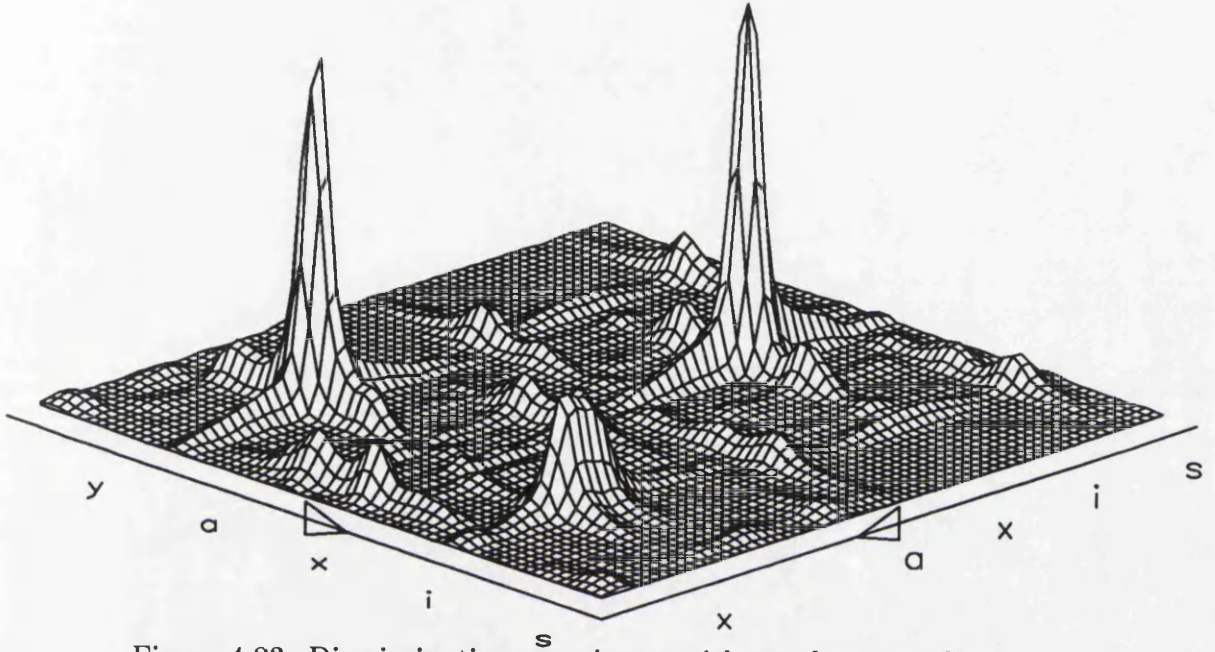


Figure 4.23: Discrimination experiment with no phase coupling.

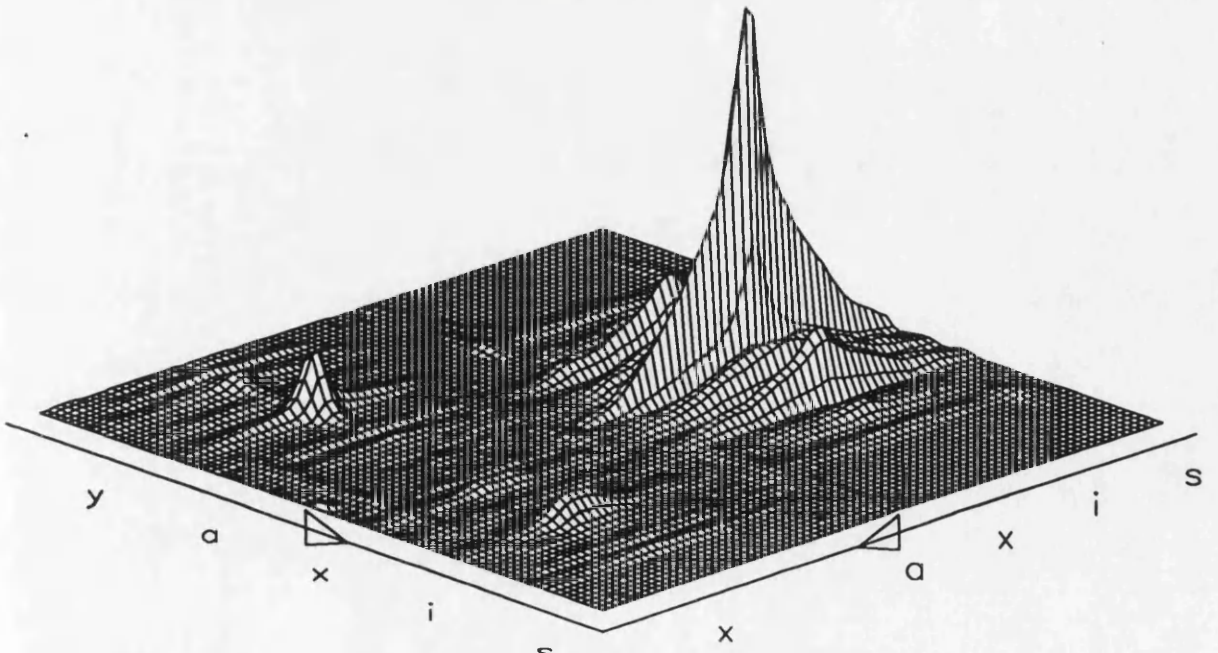


Figure 4.24: Improved discrimination as a result of phase coupling.

This is perhaps not surprising, since the unique phase pattern, associated the amplitude of an object's image, would emphasize the differences between this object and others. Such improvements in discrimination were not noted in [4.14] which merely studied the effects on the correlation peak profile. Other workers have discussed the cross-coupling effects when the SLM is used in the Fourier plane of a system [4.15] [4.16]. However, this was beyond the scope of this research, since the correlator was to have used a holographic material to store the MSFs.

4.1.7 Device Resolution

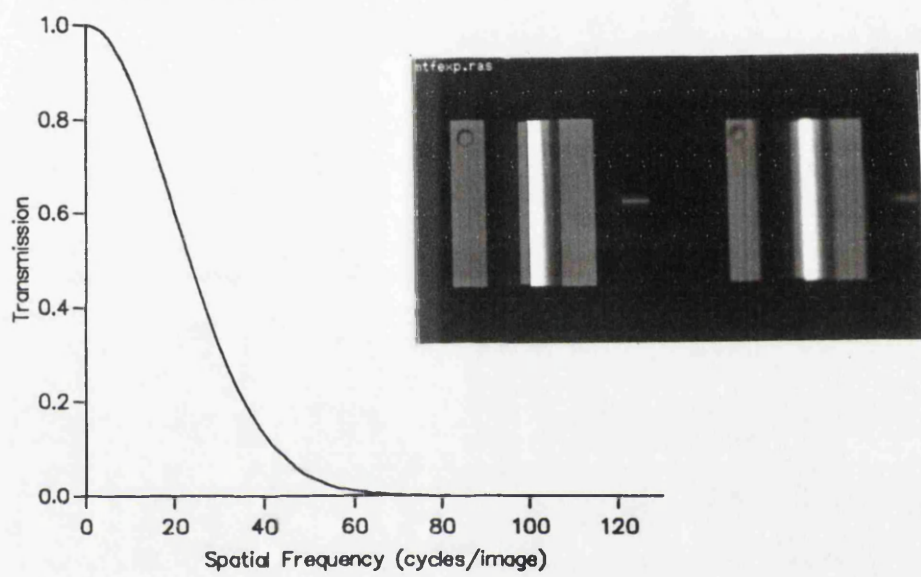


Figure 4.25: MTF curve, and the result of applying it to the image of component#2. The original image is shown on the left, the filtered image on the right.

For an optically addressed SLM, reductions in resolution are modelled by applying a smooth attenuation to the Fourier transform. This is the modulation transfer function (MTF) for the device. A convenient mathematical expression for generating MTF curves is:

$$F_{out}(r) = F(r)e^{-ar^2} \tag{4.15}$$

Where r is the radial distance from the d.c. of the Fourier transform. The shape of this curve is fairly typical of reported MTF measurements. An example curve is shown on the left of figure 4.25; the result of applying it to a 256x256 pixel image of component#2 is also shown.

4.2 A 2^N Factorial Experiment

SLMs will usually suffer from several distortions simultaneously. It is possible that several factors acting together, will produce degradations more severe than those predicted by treating each separately. If the digital simulation work is to predict the output of an optical system, a way of investigating the interaction of the various factors must be found.

The 2^N factorial experiment is a special case of the complete factorial experiment, see [4.17] for example. Both are used when an experimental system may be influenced by several different factors simultaneously. The value that a particular factor takes in a given experiment is called the *level* of that factor. The application of a set of factors, at different levels, is called a *combination*. In the 2^N factorial experiment, each of the factors is limited to just two levels. The main reason for doing this is to limit the number of combinations; and hence the number of experiments that need to be performed.

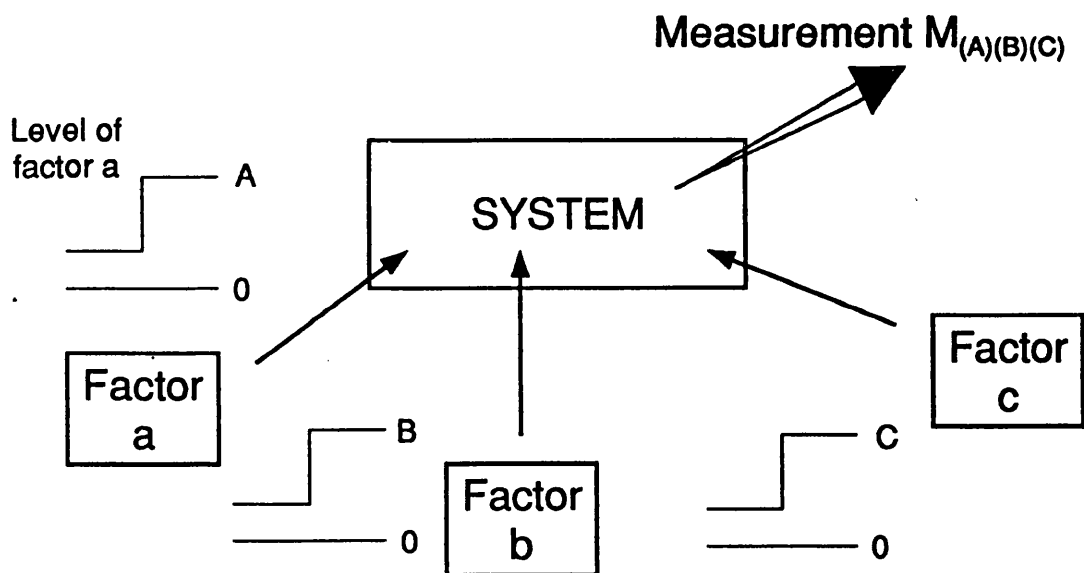


Figure 4.26: Experimental system influenced by three factors each of which can be set at two levels.

Figure 4.26 is a sketch of an experimental system which is influenced by three external factors: A , B , and C . Each of these can be present at either a 'high level' or a 'low level'. The state of the experimental system is monitored via some

measurable quantity M . To build up a picture of the effects of factors A , B and C on the value of M , experiments must be carried out using all possible combinations of the three external factors. It is usual to create an experimental nomenclature to list the results. When factor A is at a high level, and the other two are low, the measured value is denoted by M_A . Similarly, when B and C are high and A is low, the measured quantity is M_{BC} . If there are N external factors influencing the experimental system, then 2^N experiments must be performed. These experiments will cover all the possible combinations of all the factors at their low and high levels: hence the name of the experimental method.

For the system of figure 4.26, 8 experiments must be performed and eight measurements made: M , M_A , M_B , M_C , M_{AB} , M_{AC} , M_{BC} and M_{ABC} . Of these measurements, there are 4 where factor A is high and 4 where it is low. The 'direct effect' of factor A is defined by:

$$\begin{aligned} & \left(\frac{1}{4}\right) (M_{ABC} + M_{AB} + M_{AC} + M_A) \\ & - \left(\frac{1}{4}\right) (M_{BC} + M_B + M_C + M) \end{aligned} \quad (4.16)$$

i.e. the average value of M when factor A is high, minus its average value when it is low. The direct effects of the other two factors are similarly defined. Each of the eight measurements, M_{XXX} , should themselves be averaged over a number of experiments, in order to reduce the effects of random error. The direct effects can be either positive or negative, depending on whether the effect is to decrease or increase the value of M .

The main advantage of using this experimental technique, is that the interactions of the various factors can be calculated. The 'interaction' of factor A and B is defined as:

$$\begin{aligned} & \left(\frac{1}{4}\right) ((M_{ABC} + M_{AB}) - (M_{BC} + M_B)) \\ & - \left(\frac{1}{4}\right) ((M_{AC} + M_A) - (M_C + M)) \end{aligned} \quad (4.17)$$

In the first term, are grouped all the measurements for which factor B was high, and in the second term are all those for which B was low. In each of these terms, a calculation is made of the direct effect of factor A . This is done by subtracting the average value of M with A low from that with A high. Hence the 'interaction'

of A with B is the difference between the direct effect of factor A when factor B is also high and the direct effect of factor A when factor B is low. This value may be positive or negative depending whether A has a more or less positive effect on M when factor B is at a high level. It can be seen, by making the appropriate substitutions into equation 4.17, that calculating the interaction of B and A will produce a result identical to the interaction of A and B .

One problem with using the 2^N factorial experiment, is the implicit assumption that the measurable quantity varies linearly with the change in level of the factors. This problem can be alleviated by ensuring that the levels chosen are sensible in the first place and, in any case, they must be chosen to be representative of those expected to be encountered.

4.2.1 The Distortions Used in the Experiment

The images used in the experiments were those of the machined component and the hole template, shown in figure 1.3 in chapter 1. It was recognised that limits would be imposed by the number of correlations that would need to be performed in the experiment. As discussed below, it was decided to perform 12 correlations for every combination of factors. If three different distortions were included, a total of 96 correlations would need to be performed. It was felt that this would be a reasonable number if the results could be analysed automatically, without the need to take measurements from the correlation planes 'by hand'. This was feasible if the distortions were not so severe as to totally destroy the correlation plane structure. For this experiment, two features were being searched for in the image of the face of the component. Therefore, the computer scanned the correlation plane to find the two highest peaks; and then found the highest clutter signal in the rest of the correlation plane. The severity of the effects of the worst degradations was checked beforehand, by applying the maximum distortion levels to the images and performing a few sample correlations.

When selecting the three distortions to be investigated, it was noted that the input images contained no significant grey-level information. For this reason, image

dependent grey-level distortions were not included. These included dynamic range, linearity and signal dependent phase noise. Furthermore, it had already been shown that random grey-level noise had minimal effects on the correlation plane. The three distortions used were: large scale phase distortion, as described in section 4.1.4; small scale phase noise, described in section 4.1.5; and structured grey-level distortion, found in section 4.1.3. The large scale phase distortion took the form of an astigmatic phase profile with an amplitude of $\pm 2\pi$. The small scale phase distortion was generated using 3×3 pixel templates with a maximum possible distortion of 0.75π . Figure 4.27 shows the type of phase profile generated when the large and small scale phase distortions are added together. The left of the figure shows an intensity representation of the phase, ranging from black at -2π to white at $+2\pi$. On the right of the figure is a computer generated interferogram of the phase profile.

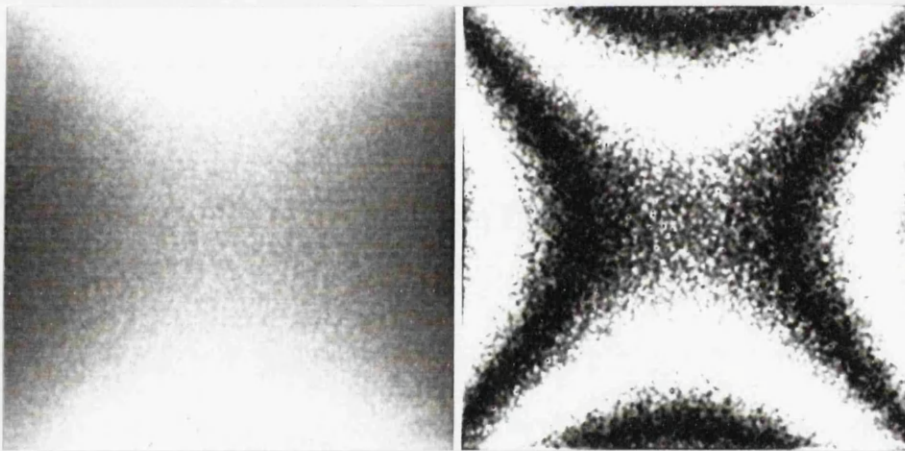


Figure 4.27: Computer generated phase profile including small scale phase distortions superimposed on an overall, large scale astigmatic distortion.

The structured grey-level distortion took the form of a non-uniform off-state. The maximum level of this distortion was such that a pixel with a grey-level of zero could be distorted to a level of 3: i.e. 1.2% of the maximum possible. The minimum level of all the above distortion factors was set to be zero; i.e. the baseline for the experiment would be a perfect SLM. To avoid correlation between the distortions on the SLMs, different patterns were applied to the object and reference images. In order to achieve this, the arrays of phase values generated were rotated by 90° between the two SLMs. Also, the function used to generate the non-uniform background used a sinusoidal

relationship for one SLM and a cosinusoid for the other (see equation 4.3).

4.2.2 Experimental Results

Four images of the machined component were captured, each rotated by 90° . This was done in order to average out any anomalies in the lighting. when the images of the component were captured. From each of these original images, two further copies were generated with the component translated towards the edge of the image. This ensured that a large area of the distortion field would be sampled when the correlations were performed. This, in turn, allowed the assessment of the tracking non-linearities as the object was translated around the field of view. There were therefore, twelve images to be correlated for every combination of distortion factors. Initial experiments suggested that the large scale phase distortion was having a relatively small effect. Subsequent calculations showed that, for all the positions of the test object in the images used, the hole features were never sampling a phase distortion greater than 0.7π . This then, was the effective phase distortion for the SLM.

Since 96 correlation planes were to be produced, it was decided to limit the size of the input images to 256×256 pixels. A computer program was written to search for the correlation peaks, due to the two holes, and then to find the peak value due to background noise or clutter. These values were all written to a file along with a description of the particular image being correlated and the levels of the three factors. Another advantage of this automated approach, was that a second program could then be written to read the data file produced and then to compile and analyse the results.

Figure 4.28 shows edge enhanced images of the component and the hole template, without any distortions present. Also shown is the correlation plane, repeated from chapter 1. As expected, the next largest signal, after that of the two holes, was produced by the round-ended slot. The signal to clutter ratio was 2:1 and 1.6:1 for the two holes. The difference in the signal strengths was due to lighting non-uniformities when the image was captured.

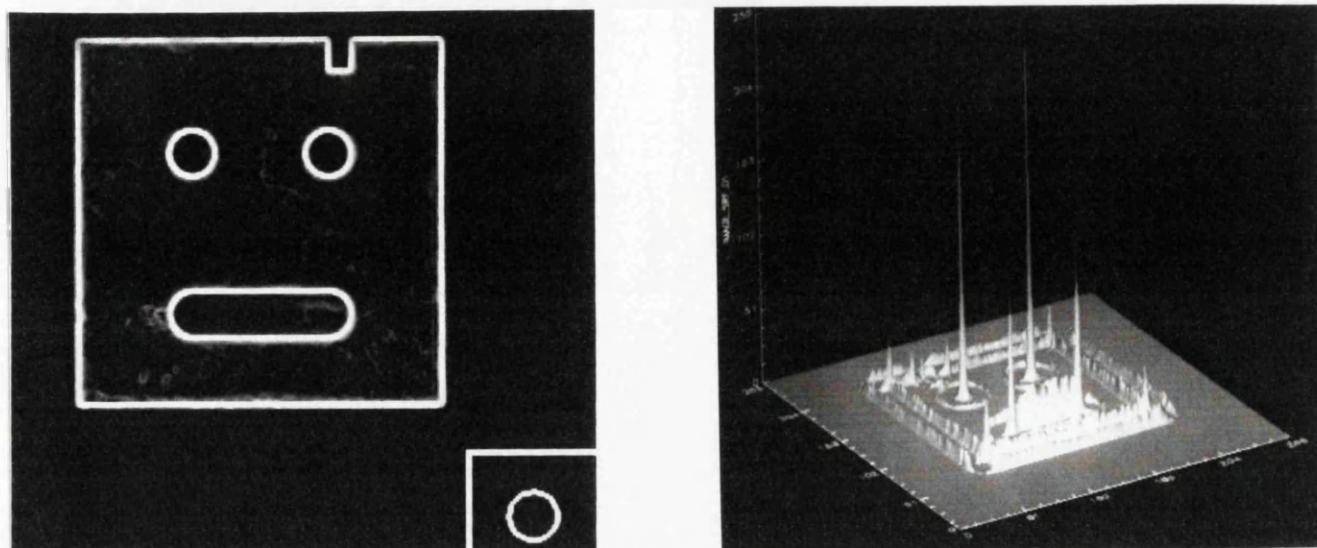


Figure 4.28: Edge enhancement of undistorted component and template, together with their correlation surface.

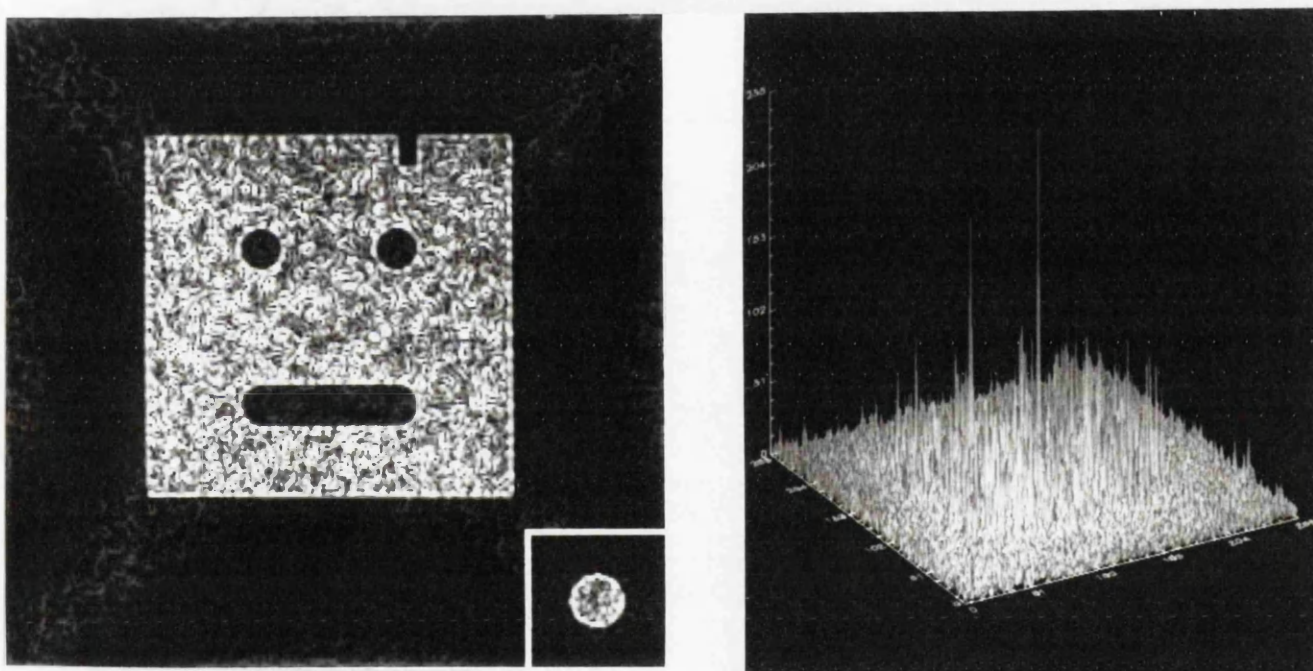


Figure 4.29: Edge enhancement and correlation results when the three distortions were present at their highest levels.

In contrast to this perfect result, figure 4.29 shows the edge enhanced images with all the distortions present at their highest levels. It also shows the corresponding correlation surface. The signal to clutter ratios had fallen to 1.28:1 and 1.19:1. The absolute intensities of the two correlation peaks corresponding to the holes – which have been normalised in the surface plots – also fell to about a quarter of their

undistorted values.

Once all the correlations had been performed, the results were collected and grouped with respect to each of the eight treatment combinations. From this information, the direct effects of the various factors were calculated. The measurable quantities used were the absolute peak intensities and the signal to clutter ratios. The discrepancies between expected and actual peak locations were also monitored. Next, the first order interactions of the factors were calculated; again using the average values of peak height and the signal to noise.

All the direct effects and interactions were calculated independently for the two holes in the image of the component. Thus, two individual estimates of these quantities were produced, which helped to give some idea of the consistency of the results. Finally, the results were searched in order to ascertain if there had been any cases in which the signal to noise ratio had fallen to below one. The results are listed below, together with the calculated values. The experimental description is written *ABC* where:

$A = 1$	\Rightarrow Large scale phase distortion	$= \pm 2\pi$
$A = 0$	\Rightarrow Large scale phase distortion	$= 0$
$B = 1$	\Rightarrow Small scale phase distortion	0.75π
$B = 0$	\Rightarrow Small scale phase distortion	$= 0$
$C = 1$	\Rightarrow Background intensity distortion	$= 3 : 255$
$C = 0$	\Rightarrow Background intensity distortion	$= 0$

DESCRIPTION: 000

Average peak hole1 = 309.6: Range: 341.1 to 288.3 = 52.9 = +/- 8.5%
Average peak hole2 = 265.5: Range: 302.1 to 233.8 = 68.4 = +/- 12.9%
Average S/C hole1 = 2.27: Range: 2.68 to 1.91 = 0.78 = +/- 17.1%
Average S/C hole2 = 1.94: Range: 2.33 to 1.64 = 0.70 = +/- 17.9%

DESCRIPTION: 100

Average peak hole1 = 269.9: Range: 325.5 to 220.7 = 104.8 = +/- 19.4%
Average peak hole2 = 230.0: Range: 288.0 to 178.4 = 109.7 = +/- 23.8%
Average S/C hole1 = 2.08: Range: 2.72 to 1.51 = 1.21 = +/- 29.0%
Average S/C hole2 = 1.77: Range: 2.47 to 1.31 = 1.16 = +/- 32.7%

Maximum position error for hole1: 0.0 pixels.

Maximum position error for hole2: 0.0 pixels.

DESCRIPTION: 010

Average peak hole1 = 206.2: Range: 279.8 to 175.8 = 104.0 = +/- 25.2%

Average peak hole2 = 188.6: Range: 232.1 to 140.4 = 91.7 = +/- 24.3%

Average S/C hole1 = 2.12: Range: 3.22 to 1.51 = 1.71 = +/- 40.4%

Average S/C hole2 = 1.93: Range: 2.71 to 1.25 = 1.46 = +/- 37.9%

Maximum position error for hole1: 1.0 pixels.

Maximum position error for hole2: 1.0 pixels.

DESCRIPTION: 001

Average peak hole1 = 160.6: Range: 193.9 to 143.9 = 50.0 = +/- 15.6%

Average peak hole2 = 135.4: Range: 150.0 to 113.4 = 36.6 = +/- 13.5%

Average S/C hole1 = 2.49: Range: 3.31 to 2.02 = 1.29 = +/- 26.0%

Average S/C hole2 = 2.09: Range: 2.49 to 1.71 = 0.78 = +/- 18.7%

Maximum position error for hole1: 0.0 pixels.

Maximum position error for hole2: 0.0 pixels.

DESCRIPTION: 110

Average peak hole1 = 185.8: Range: 256.7 to 137.5 = 119.2 = +/- 32.1%

Average peak hole2 = 163.0: Range: 211.3 to 111.9 = 99.5 = +/- 30.5%

Average S/C hole1 = 1.93: Range: 3.07 to 1.19 = 1.88 = +/- 48.8%

Average S/C hole2 = 1.68: Range: 2.45 to 1.07 = 1.38 = +/- 40.9%

Maximum position error for hole1: 1.0 pixels.

Maximum position error for hole2: 1.0 pixels.

DESCRIPTION: 101

Average peak hole1 = 111.4: Range: 133.5 to 94.8 = 38.7 = +/- 17.4%

Average peak hole2 = 93.4: Range: 114.6 to 76.0 = 38.6 = +/- 20.7%

Average S/C hole1 = 2.29: Range: 3.05 to 1.74 = 1.31 = +/- 28.6%

Average S/C hole2 = 1.91: Range: 2.47 to 1.58 = 0.88 = +/- 23.1%

Maximum position error for hole1: 0.0 pixels.

Maximum position error for hole2: 0.0 pixels.

DESCRIPTION: 011

Average peak hole1 = 113.4: Range: 163.1 to 80.7 = 82.4 = +/- 36.3%

Average peak hole2 = 105.5: Range: 138.5 to 66.4 = 72.1 = +/- 34.2%

Average S/C hole1 = 1.90: Range: 3.39 to 1.21 = 2.18 = +/- 57.2%

Average S/C hole2 = 1.74: Range: 2.13 to 1.09 = 1.05 = +/- 30.1%

Maximum position error for hole1: 1.0 pixels.

Maximum position error for hole2: 1.0 pixels.

DESCRIPTION: 111

Average peak hole1 = 81.2: Range: 117.4 to 55.0 = 62.4 = +/- 38.4%

Average peak hole2 = 72.5: Range: 103.6 to 51.6 = 52.0 = +/- 35.8%

Average S/C hole1 = 1.79: Range: 2.56 to 1.26 = 1.30 = +/- 36.4%

Average S/C hole2 = 1.60: Range: 2.35 to 1.03 = 1.32 = +/- 41.3%

Maximum position error for hole1: 1.0 pixels.

Maximum position error for hole2: 1.4 pixels.

4.2.3 Discussion of the Results

The results, for the direct effects of the three factors and their first-order interactions, are shown in table 4.1. It was noted that the two holes produced broadly similar results for the direct effects and the interactions. There were no actual failures in the recognition process. However, one of the correlations, with all the distortions present, produced a signal to clutter ratio of just 1.03:1. The positional errors produced were not large, being at most a couple of pixels, which was probably not significant. The main conclusion from the results was that the grey-level non-uniformities were by far the most damaging in terms of absolute peak intensity. They did not however, have a significant effect on the signal to clutter ratio. It must be remembered that the results are specific to this task; using the BSO correlator.

Subsequent experimentation showed that the non-uniform background was affecting the system in two ways. Firstly there was the obvious reduction in contrast of the holes, reducing the peak intensities. At the same time, the non-zero background pixels were greatly increasing the FT's d.c. intensity. This was affecting the intensity that was selected for the reference beam used to form the BSO hologram.

It was found that some of the lost correlation intensity could be recovered by not adjusting the reference beam intensity for varying d.c. levels. It was noted that the direct effects of the large scale phase distortion on peak intensity were hardly significant, especially when compared with the variations in the 'perfect' results produced by the lighting changes in the images. It should be remembered however, that for this task, the need to keep the object within the field of view, ensured that only a small part ($\sim 0.7\pi$) of the large scale phase distortion was sampled. The direct effects on signal to clutter were also barely noticeable.

Direct effect of Large scale phase on hole1 peak =	-35.39
Direct effect of Large scale phase on hole2 peak =	-34.01
Direct effect of Large scale phase on hole1 S/C =	-0.17
Direct effect of Large scale phase on hole2 S/C =	-0.18
Direct effect of Small scale phase on hole1 peak =	-66.21
Direct effect of Small scale phase on hole2 peak =	-48.67
Direct effect of Small scale phase on hole1 S/C =	-0.35
Direct effect of Small scale phase on hole2 S/C =	-0.19
Direct effect of grey-level on hole1 peak =	-126.23
Direct effect of grey-level on hole2 peak =	-110.07
Direct effect of grey-level on hole1 S/C =	0.02
Direct effect of grey-level on hole2 S/C =	0.00
Interaction large/small scale phase: hole1 peak =	9.05
Interaction large/small scale phase: hole2 peak =	4.74
Interaction large/small scale phase: S/C =	0.02
Interaction large/small scale phase: S/C =	-0.01
Interaction large phase/grey-level: hole1 peak =	-5.32
Interaction large phase/grey-level: hole2 peak =	-3.50
Interaction large phase/grey-level: S/C =	0.02
Interaction large phase/grey-level: S/C =	0.02
Interaction small phase/grey-level: hole1 peak =	27.52
Interaction small phase/grey-level: hole2 peak =	23.30
Interaction small phase/grey-level: S/C =	-0.19
Interaction small phase/grey-level: S/C =	-0.14

Table 4.1: Final results; calculated from the averaged experimental results for the different treatment combinations.

The small scale phase distortions produced a significant drop in both peak intensity and in the signal to clutter ratios. Also, the only significant interaction was between the small scale phase distortions and the grey-level non-uniformity. For the case of the correlation peak height: the positive value of this interaction implied that the *absolute* effect of the phase distortion was less severe when the grey level distortion was present. For the signal to clutter ratio, on the other hand, the negative interaction implied that the absolute effect of the phase distortion was worse with the grey-level distortion present. Thus, if the effects of the three distortions were measured separately and the results added: the predicted peak intensity would be slightly worse than that actually produced; while the prediction for the signal to clutter ratio would be on the optimistic side. To illustrate this, consider the case for the peak height produced by hole 1. The falls in absolute intensity, measured in experiments 100, 010 and 001, were -40, -103 and -149 respectively: giving a total of -292. However, the difference in peak height, measured with all the distortions present, was just -228.

The changes in signal to clutter ratio, for the same hole, for the three experiments were -0.19, -0.15 and +0.22: giving a total of -0.12. The difference in signal to clutter ratio, measured with all three distortions present, was in fact worse, at -0.48. The figures for intensity variations across the input image were calculated as a plus or minus percentage variation. It was apparent that the small scale phase distortions had the biggest effect on this variation. With all the distortions present, the percentage variation was such that a correlation peak could vary between, say, 255 and 119. This would use up a significant portion of the dynamic range of a typical CCD camera. With such a large variation it would be very difficult to do anything with the correlation plane results by way of interpretation other than simple object recognition.

4.3 Conclusions

A set of computer programs was written that could be used to apply distortions to digital images, prior to the simulation of their use in an optical processing system. The distortions, which affected the amplitudes and the phases associated with the images, were typical of those encountered when using optically addressed SLMs. Standardising the input and output formats for the programs, made it easy to apply combinations of various distortions. This was important since many different deformations will often occur simultaneously as an image is relayed by an SLM.

The small-scale phase distortion, modelled in this chapter, has not received much attention in the surveyed literature. However, it was shown to introduce a large amount of noise into edge enhanced images and correlation planes. When modelling large-scale phase errors, an important conclusion was drawn regarding the type of phase distortion that could be studied digitally. It was shown that modelling either spherical or ramp phase distortions was not valid. This was because, in an optical system, they would be compensated for by translating the Fourier plane filter. This has not been noted in previously published papers.

Another factor, not covered extensively elsewhere, was the effect of background non-uniformities in an image. It was shown that, even at low levels, they could have a serious effect on the tracking uniformity of a correlator. A study of signal-dependent phase modulation demonstrated that this effect could actually improve the discriminating capability of the correlation process.

The ultimate aim of this simulation work, was to produce software that could be used to assess the effects of various SLM distortions for any given image processing task. Since many different factors act simultaneously on the SLM output, the 2^N factorial method was used as the basis for the experiments. A real image of a machined component was correlated with a computer generated template. The task was to identify and locate two drilled holes in the component's face. Three distortions were applied to the images in various combinations, these were: large-scale phase distortion; small-scale phase noise; and structured grey-level distortion. For every

possible combination, 12 correlations were performed, with the machined component at 12 different locations in the input image. Since each distortion could be either 'on' or 'off', there were eight possible combinations of the factors and thus a total of 96 correlations was performed. One of the key steps in speeding up the experiment was the automatic analysis of the correlation planes. This was possible since: firstly, the distortions were never so bad as to completely destroy the correlation peaks; and secondly, the expected location of the peaks was known. A master program then performed the 96 correlations, tabulated the results and calculated the direct effects and interactions of the various distortions.

This method could be applied to any optical processing operation for which an SLM is used as an input. However, there must be a measurable quantity that can be defined, and that the computer can readily monitor. In the example experiment performed here, the quantities were: correlation peak height, signal to clutter ratio, and peak location accuracy. If these quantities had been measured by a human observer, for all of the 96 correlation planes, the experiment would obviously have taken much longer.

With images of 256×256 pixels, the time taken to perform 96 correlations, using a Sun SPARC station 2, was about half an hour. Using 512×512 images, the experiment would have taken four times as long. If desired, six different factors could be investigated, using 512×512 pixel images, in about 13 hours of run-time. This assumes that, for each of the 64 possible combinations of the 6 factors, the results are averaged over 12 correlations. Thus, even without specialised hardware, this experimental technique can be extended to larger images with more image distortions.

References: Chapter 4

- 4.1 Nicol, D.G., Hildebrandt, J.W., "The Effects of Random Noise on Optically Determined Auto-Correlation and Chord Functions", *Opt.Comm.*, **56** p.11 (1985).
- 4.2 Horner, J.L., "Is Phase Correction Required in SLM Based Optical Correlators?", *Appl.Opt.*, **27** p.436 (1988).
- 4.3 Vijaya Kumar, B.V.K, Cassasent, D., "Binarization Effects in a Correlator with Noisy Input Data", *Appl.Opt.*, **20** p.1433 (1981).
- 4.4 Horner, J.L., Gianno, P.D. "Phase only Matched Filtering", *Appl.Opt.*, **23** p.812 (1984).
- 4.5 Sayyah, K., "High Performance Single Crystal Silicon Liquid Crystal Light Valve with Good Image Uniformity", *Appl.Opt.*, **22** p.4748 (1983).
- 4.6 Cassasent, D., Xia, S-F, "Phase Correction of Spatial Light Modulators", *Opt.Lett.*, **11** p.398 (1986).
- 4.7 Mok, F., Diep, J., Liu, H.K., Psaltis, D., "Real-Time Computer Generated Hologram by Means of Liquid Crystal Television Spatial Light Modulator", *Opt.Lett.*, **11** p.748 (1986).
- 4.8 Juday, R.D., Daiuto, B.J., "Relaxation Method of Compensation in an Optical Correlator", *Opt.Eng.*, **26** p.1094 (1987).
- 4.9 Kim, H.M., Jeong, J.W., Kang, M.H., Jeong, S.I., "Phase Correction of a Spatial Light Modulator Displaying a Binary Phase Only Filter", *Appl.Opt.*, **27** p.4167 (1988).
- 4.10 Slack, T.G., Parker, S.C.J., "The use of Coherent Optical Processing for the Detection of Faults on a PCB", BAe Report, JS11159.
- 4.11 "Optical Information Processing", S.H.Lee (Ed), Topics in Applied Physics, Springer-Verlag 1981.

- 4.12 Haggerty, J., Young, M., "Spatial Light Modulator for Texture Classification", Appl.Opt., **28** p.4992 (1989).
- 4.13 Konforti, N., Marom, E., Wu, S-T, "Phase Only Modulation with Twisted Nematic Liquid Crystal Spatial Light Modulators", Opt.Lett., **13** p.251 (1988).
- 4.14 Horner, J.L., Giano, P.D., "Signal Dependent Phase Distortion in Optical Correlators", Appl.Opt., **26** p.2484 (1987).
- 4.15 Juday, R.D., "Correlation With a Spatial Light Modulator Having Phase and Amplitude Cross Coupling", Appl.Opt., **28** p.4865 (1989).
- 4.16 Kaura, M.A., Rhodes, W.T., "Optical Correlator Performance Using a Phase With Constrained Magnitude Complex Spatial Filter", Appl.Opt., **29** p.2587 (1990).
- 4.17 "Statistics for technology. A course in applied statistics. Third Edition", Chatfield, C., Chapman and Hall 1983.

Chapter 5

The Design of the Optical Correlator

Contents of Chapter 5

- 5.1 Calculations for the Specified Optical System 186
 - 5.1.1 Steps used in the Definition 188
 - 5.1.2 Calculated values for the System 194
- 5.2 Reducing the Set of Variable Parameters 198
- 5.3 Extending the Correlator to Larger Images 204
 - 5.3.1 Maximum Image Size on the Read-Side 205
 - 5.3.2 Large Image on the Write-Side 207
- 5.4 Conclusions 211
 - 5.4.1 Discussion of the Optical System Design 212
 - 5.4.2 Discussion of the Processing of Larger Images 214

Abbreviations/symbols used in Chapter 5

Ar^+	Argon ion (laser) (Chapt. 1).
BSO	Bismuth Silicon Oxide (Chapt. 1).
CCD	Charge Coupled Device (Chapt. 1).
CRT	Cathode Ray Tube (Chapt. 1).
d.c.	Zero-spatial-frequency component of FT (Chapt 1.).
FT	Fourier Transform (Chapt 1.).
He-Ne	Helium-Neon (laser) (Chapt. 1).
SLM	Spatial Light Modulator (Chapt. 1).
TV	Television (Chapt. 1).
B	Bragg angle for hologram (Chapt. 2).
ΔB	Miss-match between read-beam angle and Bragg angle (Chapt. 2).
d	Hologram thickness (Chapt. 2).
E_A	Electric field applied to BSO (Chapt. 2).
f	Lens' focal length.
I_{FT}	Intensity in the Fourier plane (Chapt. 3).
n	Refractive index (of BSO) (Chapt 2.)
Δn	Refractive index modulation (of hologram) (Chapt 2.).
$N \times N$	Number of pixels in an image (Chapt. 1).
P	Laser power.
Q	Hologram thickness parameter (Chapt. 2).

r_{BSO}	Half-width of clear aperture of BSO crystal (Chapt.3).
r_{corr}	Half-width of correlation plane.
r_{im}	Half-width of input image aperture (Chapt. 3).
$r_{eff} = r_{41}$	Relevant effective electro-optic coefficient of BSO (Chapt. 2).
u	Spatial frequency (Chapt. 3).
Δz	Represents the longitudinal extent of the Fourier transform (Chapt. 2).
α	Half-angle between holographic writing beams (Chapt. 2).
$\Delta\alpha$	Difference between actual α and that required to produce fringe spacing Λ (Chapt. 2).
η	Diffraction efficiency (Chapt. 2).
Λ	Hologram's fringe spacing (Chapt. 2).
Λ_C	Critical fringe spacing for BSO (Chapt. 2).
λ	Free-space wavelength (Chapt. 2).
λ_r, λ_w	Subscripts to denote reading and writing of hologram.
Θ^2	Diffusion length parameter (Chapt.3).
τ_c	Dielectric lifetime for BSO (Chapt.3).
τ_g	Exponential rate coefficient for growth of field in BSO (Chapt. 3).

The first section in this chapter, describes the system that was specified to meet the requirements of the research project. The design was produced within the constraints imposed by the available equipment and materials. The second section is concerned with the generalisation of those equations which were used to predict the system's performance. It is shown how the number of variables can be reduced to a small set which includes the number of pixels in the input image and the physical size of the BSO crystal. By examining the behaviour of these equations, predictions were made about the feasibility of extending the performance of the correlator. In the third section, conclusions are drawn about the suitability of BSO, for use as an updatable holographic medium in an optical correlator.

5.1 Calculations for the Specified Optical System

The starting point, for the design of the optical correlator, was a list of the equipment and material specifications. Some of these were fixed; being determined by the available hardware. Ideally, the dimensions of the BSO crystal would have been specified during the design process. However, the purchase and characterisation of a single BSO crystal was all that was allowed within the limits of the available budget. Therefore, correlator design is based around this specific crystal. A list of the various system parameters is given in table 5.1. Those listed as having 'variable' values, were free to be adjusted in the design stage. Those that were fixed, have their appropriate values listed. The input images, the BSO crystal, and the correlation plane detector, were all square in shape. When a radius is referred to, this is the half-width of the square. The radius of the correlation plane was based on the pickup area of a COHU 4710 CCD camera; which is a rectangle of $6.4 \times 4.8\text{mm}$.

The CCD camera sensitivity was measured experimentally. An expanded, He-Ne laser beam was used to illuminate a small aperture. This aperture was then imaged onto the CCD detector of the camera, using a long focal length lens to allow easy access. Neutral density filters were placed in the laser beam, in order to reduce the intensity to a level that did not saturate the camera.

Description	Symbol	Value
Radius of input image:	r_{im}	5.64 mm
Number of pixels across input image:	N	512
Max. spatial frequency:	u_{max}	22.7 lpmm ⁻¹
Radius of correlation plane CCD:	r_{corr}	2.4 mm
CCD camera sensitivity:	—	5nW mm ⁻²
Read laser wavelength:	λ_r	633 nm
Write laser wavelength:	λ_w	514 nm
Read laser power:	P_r	15 mW
Write laser power:	P_w	1000 mW
Refractive index of BSO:	n	2.5
Effective electro-optic coefficient:	$r_{eff} = r_{41}$	3.6 pm/V
Diffusion length parameter:	Θ^2	140
Dielectric lifetime (at 30 μ Wmm ⁻²):	τ_c	5.1ms
Required response time:	$2.7 \times \tau_g'$	40 ms
Radius of BSO:	r_{BSO}	6.5 mm
Thickness of BSO:	d	1.0 mm
Hologram grating spacing:	Λ	Variable
Critical grating spacing:	Λ_C	0.5 μ m
Read lens focal length:	f_r	Variable
Write lens focal length:	f_w	Variable
Applied field:	E_A	Variable

Table 5.1: Parameters specified for the Optical System

The camera was then replaced by a power meter and the filters removed to produce a detectable signal. This allowed the measurement of the intensity that had been needed to produce an upper-range grey-level value in the image produced by the camera.

5.1.1 Steps used in the Definition

The steps used to define the correlator's architecture are illustrated in figure 5.1. The first parameters that were set were the focal lengths of the lenses on the write and read-sides of the correlator. These were determined by: the maximum image spatial frequency; the wavelengths; and the radius of the BSO crystal. The values were calculated by substituting the relevant data into equation 3.15, of chapter 3.

The value of read-side focal length, together with the wavelength and the image size, dictated the grating spacing required for the BSO hologram. The maximum allowable spacing was calculated from equation 3.50 of chapter 3. As discussed in previous chapters, this maximum spacing would normally be used since it is beneficial to the correlator performance; both in terms of minimising angular selectivity, and in terms of maximising the speed of response.

The lens' focal length on the write-side, was one of the factors that determined the expected intensities at the BSO hologram. In order to calculate these intensities, the procedure outlined in section 3.3 was followed. The first step, involved the computation of the Fourier transforms of the two input images which were shown in figure 1.3, in the introductory chapter. Once this was done, the intensities were calculated by inserting the appropriate values into equation 3.42 of chapter 3. As well as the computed amplitude of the FT, the equation also depended on: the power of the laser; the lens focal length; and the image size.

In this calculation, account was taken of the need for the writing laser to provide both an object beam and a holographic reference beam. Assuming that there was no loss of power at the beam division stage, $P_{ob} + P_{ref} = P_{las}$; where P indicates the power, in Watts, of the various beams.

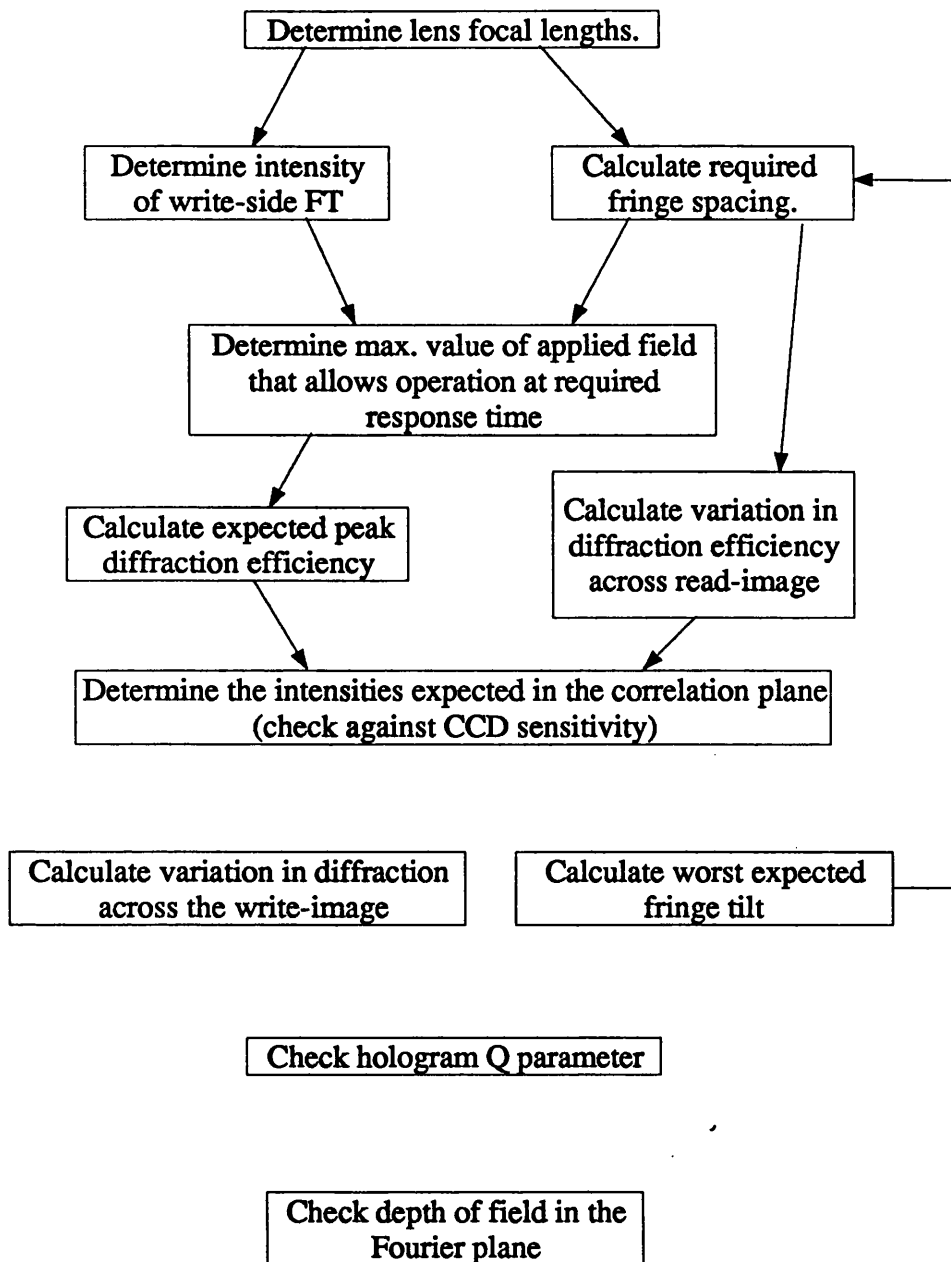


Figure 5.1: Steps used in defining the optical correlator parameters.

After the beam expansion and Fourier transforming stages, the intensity of the FT – at a specific frequency – was proportional to P_{ob} . Furthermore, after expanding the beam to flood the BSO crystal, the reference beam intensity was proportional to P_{ref} . These constants of proportionality, for the object and reference beams, can be written as O and R respectively. For the formation of the BSO hologram, the object and reference beam intensities should be equal at the spatial frequency of interest:

i.e. $OP_{ob} = RP_{ref}$ and thus:

$$P_{ob} = P_{las} \left(1 + \frac{O}{R}\right)^{-1} \quad (5.1)$$

The constant O can be found from equation 3.42:

$$O = \frac{g_{FT}^2}{255} \frac{0.05}{2\pi r_{im}^2} \frac{\beta^4}{(f_w \lambda_w)^2} \quad (5.2)$$

In the above equation, g_{FT} is the amplitude of a point in the FT; calculated by digital simulation. The constant β is the scale, in mm/pixel, of the images used in the simulation. The figure 0.05, was included on the assumption that the intensity of the expanded laser beam does not fall by more than 5% across the image. The constant R was obtained by considering the expansion of the reference beam to flood the BSO crystal:

$$R = \frac{0.05}{2\pi r_{BSO}^2} \quad (5.3)$$

The intensities in the Fourier plane were then calculated, using the value of P_{ob} in equation 3.42. Thus, the reduction in available laser power, due to the need to provide the holographic reference beam, was accounted for.

In the previous two stages, the holographic grating spacing and the writing intensities were calculated. These values, together with the measured values for Θ^2 and τ_c , were used to evaluate equation 3.46; found in chapter 3. This allowed the determination of the hologram's rise-time constant, τ_g , for any value of applied field, E_A . The processing speed was one of the predefined targets for the demonstrator. Thus, since the crystal's response slows down as E_A is increased, the maximum allowed value of applied field could be calculated. This was done assuming that the response time was the time taken to reach 90% of the maximum space charge field, $3\tau_g$. This only accounted for the grating rise-time; but an intense erase beam, derived from any light source, could be used to produce almost instantaneous decay.

By using this maximised value of E_A , the peak diffraction efficiency, η_{max} , was calculated from equation 2.11, of chapter 2. This value also relied on the expected value of Δn , which was calculated from equation 2.5 using the measured value of r_{41} . As well as η_{max} , the fall in η across the read-image was calculated. In order to do

this, the maximum Bragg miss-match was determined from equation 2.29. The drop in η , corresponding to this miss-match, was then calculated from equation 2.1. All angles were multiplied by $1/n$, to account for the refraction of the beams on their entry into the BSO crystal. For all the calculations involving the angular spread of the beams, it was assumed that the images were located close to the front focal planes of the lenses. This was necessary in order to avoid the chance of introducing complicated spatial filtering effects, as discussed in chapter 3.

The calculated value of η was used to determine the intensity of the correlation peaks in the output plane, as outlined in section 3.3 of chapter 3. In order to do this, a digital auto-correlation was carried out using the images of the machined component and the hole template. The reasons for using the auto-correlations are discussed later. The computed correlation peak values were then used, along with the relevant system parameters, in equation 3.34 of chapter 3. It was important to demonstrate that the intensities would be significantly larger than the minimum intensity that can be detected by the CCD camera. To understand what is meant by 'significantly' in this case, the experimental results of chapter 4 should be considered. The digital simulation, carried out in this chapter, computed the effects of SLM distortions on the output correlation plane intensity. Two types of phase distortion and one type of grey-level non-uniformity were modelled. The values used for the magnitudes of the distortions were considered to represent the levels expected for a real SLM. The results gave an average drop in correlation plane intensity, in the simultaneous presence of all the distortions, of about 60%, with a maximum fall of nearly 80% recorded in one of the experiments. Thus, the intensities in the correlation plane should be sufficiently greater than the camera's sensitivity to allow for these sort of variations.

At this point, all the relevant system parameters were preliminarily defined. As discussed in chapter 2, the variation in η due to hologram selectivity was expected to be relatively insignificant for the write-image. The actual variation was calculated by assessing the maximum Bragg miss-match, using equations 2.29 and 2.23. A more severe effect was expected to result from the tilting of the holographic grating vector

with respect to the direction of the applied field. The maximum fringe tilt expected in the correlator was calculated using equation 3.49 of chapter 3. This tilt was then related to a fall in η by referring to the experimental plot shown in figure 3.20.

If it was thought that the maximum fringe tilt was too large, the holographic fringe spacing would have to be decreased. This would alter the response time and the angular selectivity of the hologram, and those parameters that were affected would have to be recalculated. The only other alternative to a change in grating spacing, would be a reduction in the physical size of the write-image. This size was pre-determined by the CRT purchased for the project; which implied that the only way to reduce it was to reduce the number of pixels in the image.

The final two values calculated were: the hologram's Q parameter, and the depth of field in the Fourier plane. The Q parameter gave an indication of the extent to which the BSO would behave as a thick hologram. In order to perform the calculation, the relevant values were used in equation 2.30 of chapter 2. For the calculation of the depth of field, equation 2.39 was used.

Calculating the auto-correlation of the hole template, gave the same result as that which would be expected from a cross-correlation of the template and the image of the component. In order to perform the cross-correlation, an edge enhanced version of the hole would be stored as a Fourier transform hologram in the BSO. The edge enhancement would be achieved by setting the intensity of the reference beam to be $1/16$ that of the FT's d.c. (see later). The hologram would then be read out with the FT of the machined component. However, the only parts of this image that would contribute to the two main correlation plane peaks, would be the two holes in the face of the component. The fact that the holes, in the image of the component, are of opposite contrast to the hole template, only affects the phase of the correlation peaks. The actual intensities are the same as if the positive contrast template was being correlated with itself. This is only true because edge enhancement is being performed.

The second case to consider, is that in which the BSO hologram is formed from the Fourier transform of the image of the machined component. In order to form a

good correlation with the hole template, it would again be necessary to edge enhance the hole features. However, because the holes are only a small part of the image of the component, the intensities associated with them are relatively small. The computed d.c. intensity of the Fourier transform of the machined component was about 4,500 times that of the hole alone. Thus, in order to edge enhance the image of the hole, the reference beam would have to be set to an intensity of $1/(16 \times 4,500)$ of that of the d.c.

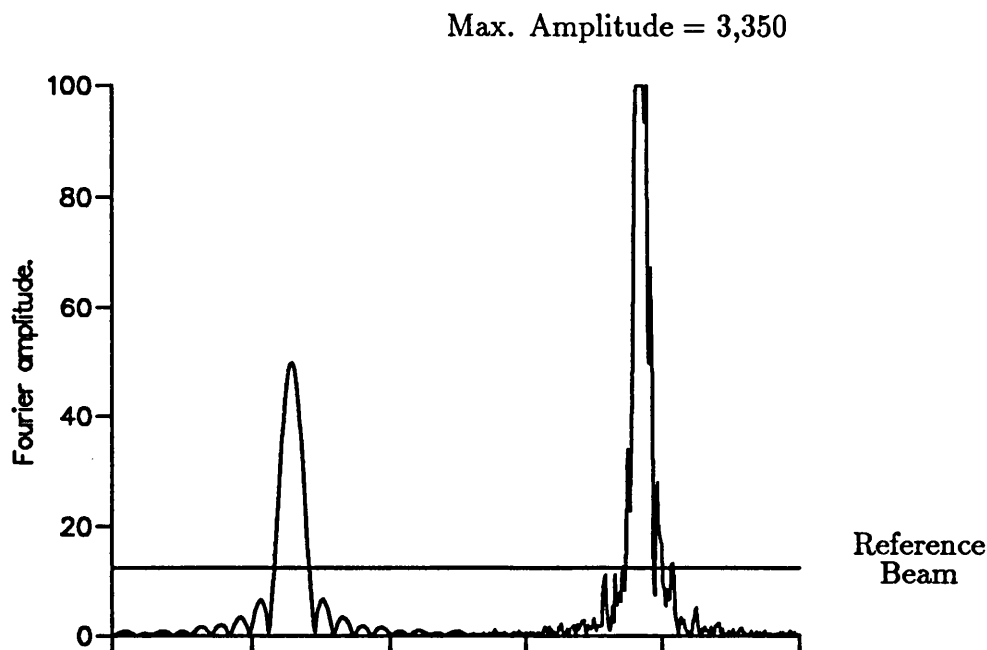


Figure 5.2: Comparison between the Fourier transforms of the hole template (left) and the machined component (right).

Two plots, taken through the centres of the Fourier transforms of the hole template and the machined component, are shown in figure 5.2. In this case, the amplitude of the Fourier transforms, which is taken to be the square root of the intensity, has been plotted. It was not possible to fit the full transform of the machined component onto the same scale as the hole template without a great loss of detail. Therefore, the values have been clipped. The amplitude of the reference beam, required to edge enhance the holes, is also shown in the figure.

This implied that, in order to perform the cross-correlation, the BSO hologram had to be formed from the Fourier transform of the hole template. If an attempt

was made to record the transform of the machined component, the reference beam intensity would be very much lower than that of the transform's d.c. In this case it was highly likely that a lot of noise, due to scattered light etc., would also be recorded. Furthermore, the very bright d.c. would perturb the electric field applied to the BSO; by creating a local region of relatively low electrical resistance. This would have effects on the correlation plane that would be difficult to predict. Local changes in electric field strength would produce local variations in the diffraction efficiency of the Fourier hologram. These, in turn, would produce some form of spatial filtering.

5.1.2 Calculated values for the System

The write and read-side focal lengths, and the maximum fringe spacing are listed below:

$$\begin{aligned} f_w &= 557.1 \text{ mm} \\ f_r &= 452.4 \text{ mm} \\ \Lambda &= 16.9 \text{ } \mu\text{m}. \end{aligned}$$

The value of the fringe spacing was significantly larger than the critical fringe spacing of $0.5\mu\text{m}$. Thus, it was expected that the space charge field in the crystal would be approximately equal to the applied field, as discussed in chapters 2 and 3.

In order to produce the desired correlations, it was required that the stored BSO holograms represented edge enhanced versions of the input images. To achieve this, the reference beam intensity had to be balanced to that of the Fourier transform at some spatial frequency higher than the d.c. Experiments were carried out with the images of figure 1.3, shown in chapter 1, using digital simulation of the BSO holograms. A good edge enhancement of the hole template was produced by making the reference beam $1/16$ of the intensity of the Fourier transform's d.c. component. For the image of the machined component, the reference beam was set at $1/256$ of the d.c. intensity to produce the required edge enhancement.

The appropriate calculations were performed for the images of the hole and the machined component. The results, assuming 1 Watt of input laser power, are shown

in table 5.2.

Image	$O(\text{mm}^{-2})$	$R(\text{mm}^{-2})$	$P_{ob}(\text{W})$	$P_{ref}(\text{W})$	$I_{FT}(\text{Wmm}^{-2})$
Hole	5.36×10^{-5}	1.88×10^{-4}	0.78	0.22	4.2×10^{-5}
Machined component	1.50×10^{-2}	1.88×10^{-4}	0.01	0.99	1.9×10^{-4}

Table 5.2: Parameters determining the writing intensities in the Fourier plane.

Given these values, the strengths of the electric fields that could be applied, while meeting the response time requirements, are as shown in table 5.3.

Image	$E_A(\text{kVcm}^{-1})$	$\tau_g(\text{ms})$	response time (ms)
Hole	2.1	13.33	40
Machined component	5.2	13.33	40

Table 5.3: Applied fields and associated response times.

For the specified size of input images – and the focal length of the lenses for this system – the maximum Bragg miss-match, on the read-side of the correlator, was 4.91mrad. For the applied field strengths calculated above, the diffraction efficiency parameters are shown in table 5.4. In the table: Δn is the refractive index modulation of the hologram, and the value for η_{max} took into account the absorption of the read-beam and the Fresnel reflections at the surfaces.

Image	Δn	η_{max}	η_{min}/η_{max}
Hole	6.1×10^{-6}	0.09%	0.75
Machined component	1.5×10^{-5}	0.58%	0.75

Table 5.4: Diffraction efficiency parameters for holograms of the two images.

Once the values for the diffraction efficiency were calculated, a digital simulation was used to determine the expected intensities in the correlation plane. In each case, the auto-correlation of the images was formed. The results are shown in table 5.5.

Image	Correlation plane intensity (nWmm ⁻²)
Hole	40.3
Machined component	526.0

Table 5.5: Correlation plane intensities for the auto-correlations of the two images.

The maximum Bragg miss-match, on the write-side of the correlator, was found to be just 0.57mrad. This would have little or no effect on the correlation output. The *Q* parameter, calculated for the BSO hologram, was 5.5; which suggested that it would exhibit thick behaviour.

The maximum fringe tilt was found to be 26.6°. From the experimental results shown in figure 3.20 of chapter 3, this would cause a fall in diffraction efficiency to between 50 and 60% of the peak value.

A diagram of the proposed optical correlator is shown in figure 5.3. The total optical path length was about 2.5 meters. This was relatively large; but was a result of the need for long focal length lenses and the restrictions on the positions of the input images.

The two SLM/CRT combinations were addressed by passing an expanded, collimated laser beam through polarising beam splitters: marked p.b.s. in the figure. The half wave plate in the read-beam controlled the intensity of the light reaching the correlation plane. That on the write-side, controlled the ratio of the intensity of the holographic reference beam to that of the object beam. The 452mm focal length lens on the read-side comprised two 600mm achromatic doublet lenses, separated by 403.5mm. The 557mm focal length lens, on the write-side, was created by placing two 750mm lenses, a distance of 490mm apart.

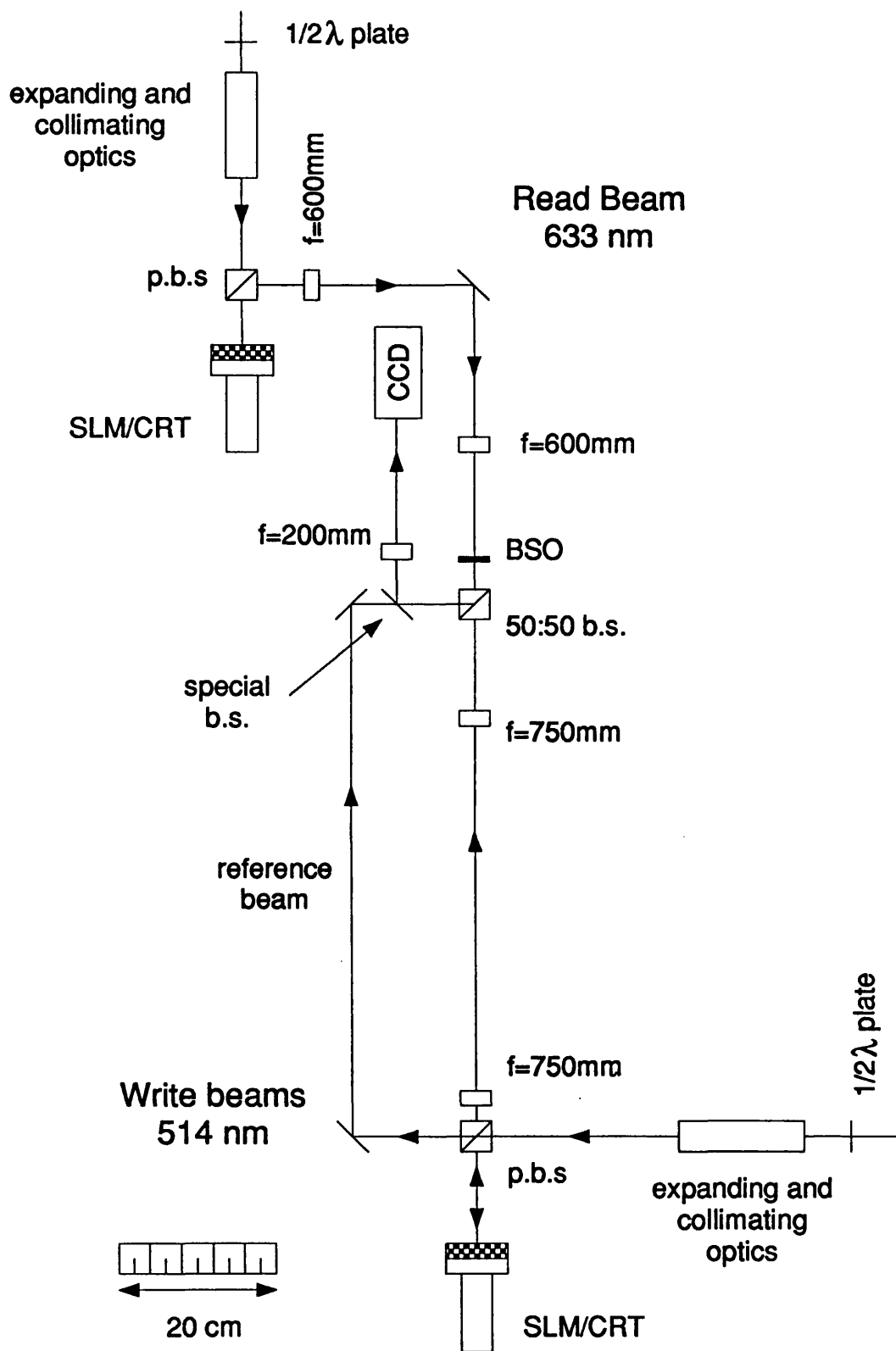


Figure 5.3: Scale diagram of the optical correlation system.

The 50:50 beam splitter, on the writing side of the BSO, performed two functions. Firstly, it was used to re-combine the holographic reference and object beams. Because of the large grating spacing used, the angle between these beams was very small. For this reason, the reference beam would not have cleared the 750mm lens before the BSO; hence the need for the beam splitter. The second function performed by the beam splitter was to re-direct the diffracted light from the read beam towards the 'special' beam splitter. This was a specially coated component which, at a 45° angle of incidence, transmitted about 90% of the argon ion beam and reflected 90% of the He-Ne read-beam.

The correlation plane was imaged onto the CCD camera using a 200mm achromatic doublet. The effects of using a cube-shaped beam splitter, placed between the lens and the Fourier plane were evaluated using computer ray tracing programs. These showed that a cube of glass introduced less aberrations than a flat plate; this latter component introduced coma into the focal spots. The only other alternative would have been to use a pellicle beam splitter. Although this component would not have introduced any aberrations, pellicle splitters are very prone to being excited by air vibrations. For this reason, they are impractical for use in interferometric systems. Not shown in the diagram are a wavelength selective filter and polariser, which would be located in front of the CCD to improve the signal to noise ratio in the correlation plane.

5.2 Reducing the Set of Variable Parameters

The BSO-based correlator that had been designed for the research project, was specified to operate on specific images. However, an important consideration was: could the same recording medium be used to correlate larger images; beyond the reach of digital systems? In order to answer this question, the equations that determined the system's performance were re-defined in terms of a smaller set of variables. This re-definition was carried out assuming that:

1. The lens focal length was chosen so that the highest frequency in the Fourier transform reached the edge of the BSO crystal.

2. The fringe spacing was chosen to be the maximum value that separated the correlation plane from the undiffracted beam.

In the following discussion, only the image's maximum spatial frequency will be of interest and the subscript, *max*, will be dropped.

Focal length, image size and grating spacing.

From equation 3.15 of chapter 3, we have:

$$f_{r,w} = \frac{r_{BSO}}{\lambda_{r,w}u} \quad (5.4)$$

The radius of the input images can also be expressed in terms of the number of pixels across the image and its maximum frequency:

$$r_{im} = \frac{N}{4u} \quad (5.5)$$

By substituting these expressions for r_{im} and f_r into the expression for the maximum fringe spacing, given in equation 3.50 of chapter 3, it is found that:

$$\Lambda = \frac{4}{3} \left(\frac{r_{BSO}}{N} \right) \quad (5.6)$$

In this case, the fringe spacing is a function of r_{BSO} and N only, all other terms have cancelled out and.

Fourier plane intensity.

The next stage in the specification of the optical system, described in section 5.1.1, was to compute the intensity in the Fourier plane. From equation 3.42 of chapter 3 it is found that, for a particular image:

$$I_{FT} \propto \frac{\beta^4}{r_{im}^2} \frac{1}{(f_w \lambda_w)^2} \quad (5.7)$$

where:

$$\beta = \frac{2r_{im}}{N} \quad (5.8)$$

As will be discussed below, it is also necessary to note that the computed values of I_{FT} are also proportional to N_{ob}^4 : where N_{ob} is the number of pixels across the actual object of interest that is present in the image.

After substituting for β , r_{im} and f_w , in equation 5.7 it is found that:

$$I_{FT} \propto \frac{N_{ob}^4}{(Nr_{BSO})^2} \quad (5.9)$$

In order to illustrate how the relationship, expressed in equation 5.9, works in practice, the following cases could be considered.

1. Increasing r_{BSO} :

This increase in r_{BSO} implies a corresponding increase in the lens focal length, which in turn increases the size of the Fourier transform. Thus, if all other factors remain unchanged, the intensity in the Fourier plane will fall.

2. Increasing N without increasing u :

Because the maximum spatial frequency, u , remains unaltered, this increase in N results in an increase in the size of the input image. Two separate examples can be looked at:

2a. N Is increased but N_{ob} remains the same:

This increase in N , results in the same sized object image embedded in a larger image field. The size of the object's Fourier transform will remain unchanged. However, the laser beam illuminating the input image will have to be expanded further to cover the larger image area. This will result in a drop in intensity at the input plane, and the intensity in the Fourier plane drops accordingly.

2b. Both N and N_{ob} are increased:

In this case, the image of the object is expanded to occupy the same proportion of the input image field. Although the intensity in the input plane will still be reduced, the larger object image will intercept the same total amount of energy. Furthermore, if the focal length of the lens is unaltered, the Fourier transform of the object image will be reduced in size. This will result in an increase in the intensities in the Fourier plane; as predicted by the equation.

3. Increasing N by increasing u :

In order to increase N , whilst leaving the physical size of the input image unaltered, the spatial frequency can be increased. In these examples, it is assumed that N_{ob} increases proportionally with N . Again, two examples can be considered.

3a. Increasing u while leaving r_{BSO} unaltered:

Because r_{BSO} remains the same, the focal length of the lens must be reduced in order to match the size of the transform to the BSO crystal. Since the physical size of the input image remains unchanged, the intensities in the input plane are unaltered. Hence, the power in the Fourier plane is increased by the greater focussing power of the lens.

3b. Increasing both u and r_{BSO} :

Since r_{BSO} is increased to accommodate the change in size of the Fourier transform, the focal length of the lens remains fixed. With the same focal length, and the same size of input images, the intensities in the Fourier plane remain unaltered.

The last two examples need some qualification. If the spatial frequency in the input image is increased, finer details may be resolved in the image of the object. These finer details will spread light to the higher frequencies in the Fourier plane. This means that predictions about intensities may give results that are slightly too high. However, as the intensities of the higher frequencies are generally very low compared to the d.c., the errors are expected to be small.

Response time.

The parameter governing the response time of the system, τ_g , depends on both I_{FT} and E_A . If I_{FT} can be increased, the improved response time could allow the system to run at faster processing rates; or the applied field could be increased while maintaining τ_g . The increase in E_A would be beneficial in increasing the diffraction efficiency of the system.

There is however, another factor which affects the apparent diffraction efficiency of the system. This is the variation in diffraction efficiency across the image.

Diffraction efficiency variation.

The maximum Bragg miss-match, on the read-side, can be found by substituting the expressions of equations 5.4 and 5.5 into equation 2.29 of chapter 2. If this is done – and the result is divided by the refractive index – the worst Bragg miss-match,

inside the crystal, is found to be:

$$\Delta B_{max} = \frac{\lambda_r}{4n} \left(\frac{N}{r_{BSO}} \right) \quad (5.10)$$

By substituting this expression into equation 2.1, in chapter 2, the value of η_{min}/η_{max} can be plotted as a function of N/r_{BSO} . Results are shown in figure 5.4, for three different hologram thicknesses.

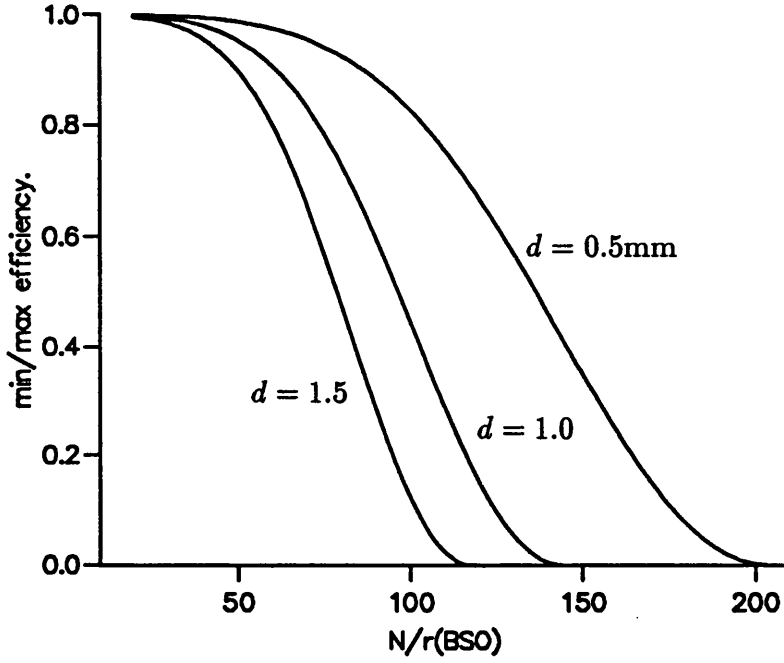


Figure 5.4: η_{min}/η_{max} plotted as a function of N/r_{BSO}

In addition to the Bragg miss-match on the read-side, that on the write-side can also be expressed in terms of N/r_{BSO} . This is done by substituting for $\Delta\alpha$ in equation 2.23, in chapter 2. The expression for $\Delta\alpha$ is identical to that for ΔB given in equation 5.10, but with the write-beam wavelength replacing that of the read beam. When the substitutions are made, the following expression is obtained:

$$\Delta B(\Delta\alpha)_{max} = \frac{(\lambda_r - \lambda_w)}{8n} \left(\frac{N}{r_{BSO}} \right) \quad (5.11)$$

Minimum hologram fringe spacing.

From equations 2.17 and 2.19, in chapter 2, the following expression can be derived:

$$\Lambda(\Delta\alpha) = \frac{\lambda_w}{n\Delta\alpha + (\lambda_w/\Lambda)} \quad (5.12)$$

This gives the fringe spacing that results when the writing beam is moved through an angle, $\Delta\alpha$. In this case, Λ is the fringe spacing that was being produced before the writing beam was moved through the additional angle. By substituting for $\Delta\alpha$ and Λ , it is found that the minimum fringe spacing, produced in any part of the hologram, is given by:

$$\Lambda_{min} = \frac{r_{BSO}}{N} \quad (5.13)$$

Correlation plane intensities.

The intensities in the correlation plane depend on the lens focal lengths and the size of the input image as well as on the diffraction efficiency. From equation 3.34, in chapter 3, it can be seen that:

$$I_{corr} \propto \left(\frac{f_r}{r_{im}f_c} \right)^2 \quad (5.14)$$

However, the focal length f_c of the lens which forms the correlation plane, is chosen to match r_{im} to the required correlation plane size, r_{corr} . Thus:

$$\frac{f_r}{f_c} = \frac{r_{im}}{r_{corr}} \quad ; \quad (5.15)$$

and the correlation plane intensity is independent of the size of the read-image.

Hologram fringe tilt.

An expression for the maximum fringe tilt that would be present in the hologram is obtained from equation 3.49, in chapter 3. This tilt occurs for the point in the input image: $x_{im} = -r_{im}$; $y_{im} = r_{im}$. In this case, the fringe tilt is given by:

$$\cos^2(\theta) = \frac{(f_w\lambda_w - \Lambda r_{im})^2}{(f_w\lambda_w - \Lambda r_{im})^2 + (\Lambda r_{im})^2} \quad (5.16)$$

However, from equation 5.6:

$$\Lambda r_{im} = \frac{4}{3} \times \frac{r_{BSO} r_{im}}{N} \quad (5.17)$$

and by combining equations 5.4 and 5.5:

$$f_w \lambda_w = \frac{r_{BSO}}{u} = \frac{12}{3} \times \frac{r_{BSO} r_{im}}{N} \quad (5.18)$$

It can then be seen that the quantity $r_{BSO} r_{im} / 3N$ is common to all the terms in the numerator and denominator of equation 5.16. After this has been cancelled out, the expression becomes:

$$\cos^2(\theta) = \frac{(12-4)^2}{(12-4)^2 + 4^2} \Rightarrow \theta = 26.6^\circ \quad (5.19)$$

which is true for all optimised systems.

Hologram Q parameter.

The expression for Λ can be substituted into equation 2.30 in order to derive the hologram's Q parameter:

$$Q = \frac{2\pi \lambda_r d}{n} \left(\frac{N}{r_{BSO}} \right)^2 \quad (5.20)$$

Fourier plane depth of field.

The final parameter considered in this section, is the depth of field in the Fourier plane. This depends on the focal length of the lenses and on the aperture size; as can be seen from equation 2.39 in chapter 2. After making the appropriate substitutions it is found that:

$$\Delta z_{r,w} = \frac{64n}{\lambda_{r,w}} \left(\frac{r_{BSO}}{N} \right)^2 \quad (5.21)$$

5.3 Extending the Correlator to Larger Images

In order to assess the feasibility of processing larger images, many of the key system parameters have been expressed in terms of the coefficient N/r_{BSO} (units L^{-1}). Before discussing the results, it is important to understand the significance of this variable.

To increase the value of N , is to increase the number of pixels across the input image. This can be achieved in one of two ways. Firstly, the image size, r_{im} , can remain fixed and the maximum spatial frequency, u , increased. Secondly, the size of an image can be increased but its maximum spatial frequency held constant. The first option requires that either the lens focal length is decreased or that the size of the BSO crystal is increased. In the second option, the lens focal lengths will remain fixed, as will the size of the BSO crystal. Thus, in the first case, the value of N/r_{BSO} will remain fixed; whereas, in the second, it will increase.

Increasing the value of r_{BSO} allows longer focal length lenses to be used in the system. The increase in crystal size would be chosen to match the increase in the size of the Fourier transform of an image. When an increase in r_{BSO} is mentioned, for the optimised system, an increase in lens focal length is also implied.

Decreasing the value of r_{BSO} can imply one of two things. Firstly, it can mean reducing the size of the crystal and reducing the lens' focal length; so that the Fourier transform is not clipped by the BSO. Secondly, the focal length of the lens may be reduced so that the Fourier transform only occupies a part of the area of the crystal. In either case, reducing r_{BSO} implies a reduction in the lens' focal length.

5.3.1 Maximum Image Size on the Read-Side

Equation 5.13, in the previous section, expresses the minimum holographic fringe spacing that will be produced in the correlator. As was discussed in section 3.5 of chapter 3, the fringe spacings must be greater than the BSO's critical fringe spacing, Λ_C . Thus:

$$\Lambda > \Lambda_C \Rightarrow \frac{r_{BSO}}{N} > \Lambda_C \Rightarrow N < \frac{r_{BSO}}{\Lambda_C} \quad (5.22)$$

Equation 5.21, gave an expression for the depth of field in the Fourier plane. Since this should always be greater than the hologram thickness, d , another limit is imposed on N :

$$\Delta z > d \Rightarrow \frac{64nr_{BSO}^2}{\lambda_r N^2} > d \Rightarrow N < 8r_{BSO} \sqrt{\frac{n}{d\lambda_r}} \quad (5.23)$$

In addition to these, there is yet another limit imposed on N , due to the angular selectivity of the hologram. From the curves of figure 5.4, it is seen that there is a value of N/r_{BSO} at which the diffraction efficiency falls to zero for points at the edge of the field of view.

Going back to equation 2.1, of chapter 1, and noting that, at large values of Bragg miss-match $\xi^2 \gg \nu^2$ it can be seen that:

$$\frac{\eta_{min}}{\eta_{max}} = \frac{\sin^2(\sqrt{\xi^2 + \nu^2})}{\xi^2 + \nu^2} \approx \left(\frac{\sin \xi}{\xi} \right)^2 \quad (5.24)$$

This expression will equal zero when $\xi = \pi$. Thus, for the BSO correlator, the following conditions must be met:

$$\xi < \pi \Rightarrow \frac{\Delta B \pi d}{\Lambda} < \pi \quad (5.25)$$

where small angle approximations have been used in equation 2.2. Substituting for ΔB and Λ , and rearranging, gives:

$$N < r_{BSO} \sqrt{\frac{16n}{3d\lambda_r}} \quad (5.26)$$

There are now three conditions which can limit the maximum value of N on the read-side of the correlator. Using: $\lambda_r = 633\text{nm}$, $n=2.5$ and $\Lambda_C = 0.5\mu\text{m}$; the limits become:

Condition	Limit
Limit on Λ_C	$N < (2000 \times r_{BSO})$
Limit on ΔZ	$N < (500 \times r_{BSO} / \sqrt{d})$
Limit on ξ	$N < (145 \times r_{BSO} / \sqrt{d})$

Of these, the strictest limit is imposed by the Bragg miss-match term, ξ . From equation 5.20, it is seen that the hologram's Q parameter also increases with N . Thus, the BSO will be behaving more and more like a thick hologram. For the 1mm thick, 13mm diameter BSO crystal, used in the previous design work, N would be limited to 940 pixels. As discussed in the previous section, the correlation plane intensities are not expected to be affected, other than by diffraction efficiency changes.

By reducing d to 0.2mm and increasing the diameter to 26mm, N could be increased to nearly 3000 pixels. Compared to the system designed previously, the peak diffraction efficiency would be reduced by a factor of 25 times. The efficiency at the edge of the field of view might be just 1% of the Bragg-matched value at the centre of the image. Thus, with the reduction in d and the increase in the variation of diffraction efficiency across the image, correlation plane intensities could be 1875 times lower than those produced by the system specified for 512×512 pixel images. To make up for this, a pair of holograms could be used to re-profile the laser beam intensity. Even if the hologram pair had an overall efficiency of 50%, this would lead to intensities 10 times higher than those produced by expanding the beam. Efficiencies of this magnitude should be achievable with materials such as dichromated gelatin. Increasing the applied field by a factor of two, would lead to a four fold increase in intensity. These steps would be sufficient to produce intensities on the borderline of the CCD camera's sensitivity. However, this assumes that the number of pixels across both objects being correlated was similar to the images' size. If it was wished to correlate with smaller templates, then either the laser power would have to be increased or some form of image intensifier used in the correlation plane. This latter option might be ruled out by signal to noise considerations.

5.3.2 Large Image on the Write-Side

From the preceding section, it seems that it will be difficult to perform correlations with large images on the read-side of the correlator. Another possibility is to perform correlations with very large image on the write side and a small reference template on the read side. In all the prior work, it was assumed that N , the number of pixels across the image, was the same on both read and write-sides. However, there are now different values, N_r and N_w pixels; and it will be assumed that $N_w \gg N_r$. The corresponding image radii will be written r_r and r_w .

One important consequence of this change is that the size of the correlation plane will now be approximately equal to r_w . The required separation between the centre of the correlation plane and the undiffracted beam will now be $r_r + r_w \approx r_w$. This

is less than the value of $3r_{im}$ in section 3.5; where all the images were the same size. Because of this, the minimum required fringe spacing will become:

$$\Lambda = \frac{\lambda_r f_r}{r_w} = \frac{4r_{BSO}}{N_w} \quad (5.27)$$

where it has been assumed that the spatial frequencies of the write and read images are the same; although they contain different numbers of pixels. This would be beneficial in making the hologram 'thinner', and therefore less angularly selective. However, the maximum fringe tilt, produced in the hologram would now increase. Substituting the new value of Λ into equation 5.16, it is found that $\cos^2(\theta) = 0$. This is because the light from the edge of the write-image has become co-linear with the reference beam. As it will be seen later, it is, in fact, undesirable to decrease the angle between the beams.

Leaving the fringe spacing for now, the Bragg miss-match can be considered, and an equation can be written down, analogous to 5.25 above. In this case, instead of ΔB , the expression for $\Delta B(\Delta\alpha)$, from equation 5.11 is substituted. If this is done, and the equation rearranged it is found that:

$$N_w < r_{BSO} \sqrt{\frac{32n}{3d(\lambda_r - \lambda_w)}} \quad (5.28)$$

For the moment, the fringe spacing is set as if the images on the read and write sides were the same size; so equation 5.6 was used to substitute for Λ .

When the depth of field in the Fourier plane is considered, the condition is the same as that of the previous section; but with λ_w instead of λ_r . It is then found that:

$$N_w < 8r_{BSO} \sqrt{\frac{n}{d\lambda_w}} \quad (5.29)$$

When values of 633nm, 514nm and 2.5, are used for λ_r , λ_w and n respectively; the conditions on N_w become:

Condition	Limit
Limit on ΔZ	$N_w < 558 \times r_{BSO} / \sqrt{d}$
Limit on ξ	$N_w < 473 \times r_{BSO} / \sqrt{d}$

Hence, it can be seen that the strictest limit on N is still enforced by the maximum Bragg miss-match produced in the correlator.

With the values of $r_{BSO} = 6.5\text{mm}$ and $d = 1.0\text{mm}$, the system would be capable of processing write-side images of about 3000 pixels across. By reducing d to 0.3mm and increasing r_{BSO} to 9.3mm ; this figure could be increased to over 7000 pixels. For the calculations carried out below, the case of a 6000 pixel image is considered.

It now remains to calculate the diffraction efficiency variation across this large image. Because it was assumed that the read image was small, there would be no significant Bragg de-tuning on the read-side. On the write-side however, two effects would be present: the fringe tilt and the Bragg de-tuning.

Because the fringe spacing for the system was still governed by equation 5.6, equation 5.19 – previously derived for fringe tilt – still holds. This gave the angle of tilt as 26.6 degrees. In order to reduce this, the angle between the holographic object and reference beams could be increased. However, the minimum fringe spacing, produced by light from the edge of the 6000 pixel image, would already be 1.55 microns. This is getting very close to the BSO's critical fringe spacing of 0.5 microns, so there is little or no scope for reducing the fringe spacings any further. From the graph of experimental results, shown in figure 3.20, a tilt of 26 degrees would cause a drop in diffraction efficiency of 50%. This drop would occur at the two corners of the image that were nearest to the line of the reference beam.

The maximum Bragg miss-match for the write-image can be found from equation 5.11. The hologram's Q parameter, given by equation 5.20 is nearly 200; so it would definitely be expected to exhibit thick behaviour. Using the value of Bragg miss-match, and the average fringe spacing ($2\mu\text{m}$), to calculate ξ , it is found that:

$$\xi = 1.81 \Rightarrow \left(\frac{\sin \xi}{\xi} \right)^2 = \frac{\eta_{min}}{\eta_{max}} = 0.3 \quad (5.30)$$

Thus, the combined effects of Bragg de-tuning and fringe tilt, could lead to a drop in diffraction of 85% between the centre and the corners of the image.

In section 2.2.2 of chapter 2, it was stated that variations across the write-image could not be compensated for. However, when the input image on the read-side is small, there is a way round the problem.

Consider the example images shown schematically in figure 5.5.

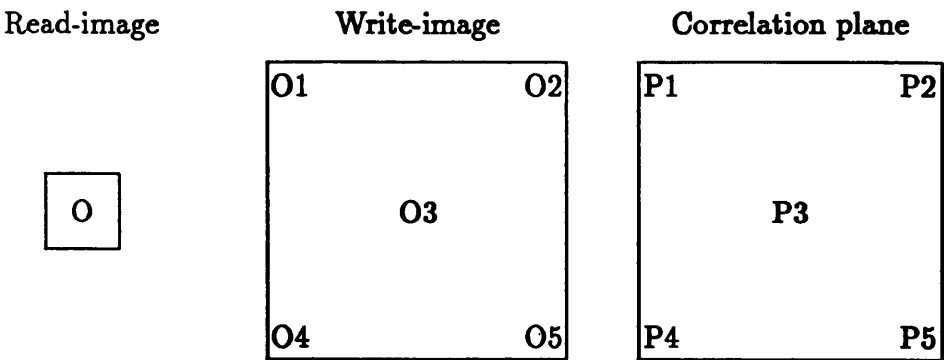


Figure 5.5: Schematic diagram of images and correlation plane for a small reference template correlated with a large image.

Because the template on the read-side is relatively small, the locations of any peaks in the output plane, match the locations of the features in the write-image. Because of fringe tilting and Bragg effects, the correlation peaks will have different intensities. The central peak will be the highest; next will be the peaks at the image corners away from the reference beam; and the peaks at the image corners nearest the reference beam will have the lowest intensities. To compensate for this, an intensity transmission filter could be placed in the correlation plane. The transmission profile of this filter would be such that the peaks would be variously attenuated according to their position. Thus, the situation is now analogous to that previously discussed for the read-side: the apparent efficiency of the hologram is reduced to the lowest value encountered for any part of the image.

This being the case, the expected drop in efficiency of 85% could easily be dealt with. Beam-shaping holographic elements could be used to improve light utilisation on the read-side, and this would compensate for some of the loss of intensity. Also, because higher intensities are expected in the write-images, higher electric fields could be applied to the BSO crystal without degrading the speed of response.

One final design feature must be incorporated for this large write-image correlator. As discussed at the end of section 5.1.1, the very intense d.c. Fourier term, from a large write-image, could seriously perturb the BSO hologram. It would therefore be necessary to include a pre-filter to block out the d.c. and, perhaps, attenuate some

of the lower spatial frequencies.

5.4 Conclusions

Firstly in this section, the broad conclusions of this chapter are stated. The individual topics of the specific system design and the extension to larger images are then discussed in more detail.

The following conclusions, regarding the practical use of BSO-based correlators, have now been drawn:

1. A correlator could be designed to meet the following requirements:

- Process images of 512×512 pixels.
- Carry out processing at TV frame rates.
- Use 15mW He-Ne laser for reading and a 1W Ar^+ laser for writing, with no special holographic elements for beam re-profiling.
- Use a standard CCD camera as the output plane detector.

2. The intensities in this system, both in the correlation plane and the Fourier plane were starting to approach the limits dictated by response time and CCD sensitivity.

3. Extending the performance of the system to cross-correlate two, 3000×3000 pixel images, might just be possible. This is very near the practical limits of producing detectable signals in the correlation plane. The effective diffraction efficiency could be as low as 0.0002%. Images of 1000–2000 pixels may represent a more feasible upper limit.

4. The angular selectivity of the BSO hologram is the main problem encountered. This is because the low refractive index modulation of the BSO necessitates the use of relatively thick crystals.

5. By making use of the reduced selectivity on the write-side, a BSO correlator could theoretically be used to correlate a 6000×6000 pixel image with a small reference template.
6. The total optical path length in the correlators will be of the order of meters, unless two very small images are being cross-correlated.
7. Extending the correlator's performance requires the fabrication of thin, large area crystals: $\sim 20 \times 20 \times 2\text{mm}$.

5.4.1 Discussion of the Optical System Design

The image processing task to be performed, involved the use of correlation to identify sub-features within a 512×512 pixel image. The identification was to take place on a time scale commensurate with TV frame rates; i.e. 40ms. The available reading and writing laser powers were limited to 15 milliWatts and 1 Watt respectively. The main conclusion, drawn from this simulation, was that the BSO correlator could be used to perform the specified correlations. Furthermore, the correlator design made use of 'off-the-shelf' optical components. However, as may be seen from table 5.5, the predicted intensities in the correlation plane approached the limits of the CCD camera's sensitivity. The reason for this was the relatively small size of the image of the hole that was used to form the BSO hologram. This led to a low writing intensity. As a result of this, only a relatively small field could be applied to the crystal; without the response time becoming unacceptably long. The output intensities may not have been sufficient if correlations were performed with SLM distortions present in the input images.

From the Fourier plane intensities of table 5.2, it is seen that there would be a definite advantage in correlating images which contain larger objects. For the image of the machined component, the writing intensity in the Fourier plane was increased by a factor of 4.5 times. This allowed a much larger electric field to be applied, while maintaining the required response time. This, in turn, increased the diffraction efficiency of the hologram. The intensity in the correlation plane was then almost 100 times greater than the sensitivity of the CCD camera. This would give

better tolerance to any reduction in the height of the correlation peak caused by SLM distortions. The angular selectivity of the BSO hologram did not, in this case, pose a serious problem. From table 5.4 it is seen that the diffraction efficiency at the edge of the read-image only fell to 75% of its maximum value at the centre. This variation would be removed by use of a transmission filter, which would reduce the apparent diffraction efficiency – for all points in the image – to 75% of the maximum. If the holograms were being written with images of large objects, there would be a problem due to the tilting of the fringes with respect to the direction of the applied field. As discussed previously, the only way to alleviate this problem, was to use smaller grating spacings. This would reduce the useable diffraction efficiency in two ways. Firstly, the angular selectivity of the hologram would increase; leading to a greater fall in diffraction efficiency across the read-image. Secondly, the response of the crystal would become slower; and the applied field would have to be reduced, in order to meet the specifications.

For the images used here, the worst-case fringe tilt could be reduced from 26.6° to 13.8° , by reducing the grating spacing to 10.0 microns. From the experimental results, this tilt caused η to fall to between 80 and 90% of its maximum value. It should be noted that the scatter of the experimental results makes an exact prediction difficult. At this spacing, the maximum field that could be applied was reduced to 3.2kV/cm. At the same time, the diffraction efficiency at the edge of the read-image fell to about 40% of the maximum value. The predicted intensity in the correlation plane then dropped to 112nW/mm^2 ; but this is still a long way above the CCD's minimum sensitivity.

Results were simulated for the hole template correlated with the machined component and for the machined component correlated with itself. The intensities of the correlation peaks, in these two cases, differed by a factor of 13 times. If the more intense peak produced a CCD output grey-level of 255, then the weaker peak would only produce a level of 20. If the system was expected to perform correlations on both of these image pairs, the range of this variation would have to be reduced. This would best be achieved by reducing the power of the read-beam when performing

the correlation which produced the greater intensity.

5.4.2 Discussion of the Processing of Larger Images

In order to assess the possibility of processing of images larger than 512×512 pixels, a number of key parameters were expressed in terms of N/r_{BSO} . N was the number of pixels across the image, and r_{BSO} was the radius of the clear aperture of the BSO crystal. The parameters expressed in terms of this new variable were:

Maximum hologram fringe spacing	(Equation 5.6)
Minimum hologram fringe spacing	(Equation 5.13)
Fourier plane Intensity	(Equation 5.9)
Read-side Bragg miss-match	(Equation 5.10)
Write-side Bragg miss-match	(Equation 5.11)
Hologram thickness parameter	(Equation 5.20)
Fourier plane depth of field	(Equation 5.21)

In addition, it was shown that the maximum hologram fringe tilt would be independent of the write-image size. Also the correlation plane intensity was independent of the read-image size.

Two cases were studied, the first involved the use of large images on the read-side, the second looked at the limits imposed on the write-side image size. The ultimate limit in the case of the read-side, was imposed by the variations in diffraction efficiency due to the angular selectivity of the hologram. The ultimate figure of 3000 pixels is about half the size of images acquired by some earth-observation satellites and by high resolution, airborne synthetic aperture radar systems. However, this figure relied on a BSO crystal of $26 \times 26 \times 0.2$ mm and it is not known, if BSO crystals of these dimensions can be produced. Furthermore, it might be difficult to create a uniform applied electric field over this size of crystal. Actually achieving this figure would depend on the light utilisation of the system. The calculated number of pixels corresponds to a correlation plane intensity at the limits of the typical sensitivity of a CCD camera. SLM distortions could reduce these intensities to levels that are too low to be practical, and light scattered from various optical surfaces could mask the correlation plane. The figure of 3000 pixels therefore represents an upper limit and a real system might only be capable of processing images of 1000 – 2000 pixels.

In the case of a large image on the write-side, the ultimate limit was imposed by the need to keep the longitudinal extent of the Fourier transform to a value greater than the thickness of the BSO crystal. Also, in order to allow the compensation for variations in diffraction efficiency across the large write-image, it was necessary to restrict the system to relatively small images on the read-side. However, it might be feasible to search a 6000×6000 pixel image for a small sub-feature using this correlator. The small reference template on the read-side of the correlator could be updated at an arbitrarily fast rate. The comments made in the previous paragraph, concerning light utilisation, would again be true here. Practical uses of such a correlator would probably be in the field of photo-reconnaissance. Airborne devices such as infra-red line-scan and synthetic aperture radars are capable of producing image data covering large swaths of terrain. Images several thousand pixels across are not uncommon. Using the correlator proposed here, these images could be rapidly searched for targets such as vehicles, ships and parked aircraft.

Chapter 6

Conclusions and Recommendations

Contents of Chapter 6

6.1	Design Meeting Required Specifications	217
6.2	Extending the Design to Larger Images	218
6.3	Digital Simulation of SLM Performance	220
6.4	Recommendations for Future Work	221

During the course of this research, advances were made in three main areas.

6.1 Design Meeting Required Specifications

A thorough optical design was undertaken and plans were produced for a BSO-based correlator. The design enabled the theoretical performance of the system to attain two specific goals: the processing of images of 512×512 ; at a rate of 40ms per frame. The system was also constrained by a need to use cw lasers with output powers of no more than 15mW and 1W for the read and write-beams respectively.

The design process incorporated several aspects that have not been discussed previously in published papers on BSO correlators. The Bragg matching constraints, imposed by the presence of a thick hologram in the Fourier plane, were investigated more thoroughly than elsewhere. The angles of light rays at the hologram were explicitly linked to the Fourier transforming process. Methods of dealing with the hologram's angular selectivity, by use of intensity transmission filters, were also discussed. It was concluded that the effects on the read-side could be negated by placing a filter, with a spatially varying intensity transmission, in the image plane. The penalty associated with this, was a reduction in the apparent diffraction efficiency to that of the minimum value which is encountered at the edges of the read-image. In the case of the write-image, action could only be taken if the correlation was being performed with a small reference template. In this case, a transmission filter could be placed in the correlation plane.

The effect of the location of the write and read-images, with respect to their transforming lenses, was also investigated. An expression was derived for the location of the correlation plane as a function of image position. However, it was also shown that complicated spatial filtering effects would result if the images were moved away from the front focal planes of the lenses. This restricted the correlator to a '4f' design.

A computer program was written that could be used to calculate the intensities at the Fourier and correlation planes of the optical system. The intensity in the Fourier plane governed the response time of the system for a given magnitude of electric field

applied to the BSO crystal. As the electric field is increased, there is an associated rise in diffraction efficiency does; but the speed of response falls. Thus, for a given Fourier plane intensity, there will be a maximum field that can be applied, whilst maintaining the required system response time. This limit imposed on the strength of the applied field will, in turn, limit the diffraction efficiency of the BSO hologram.

Lens design work was undertaken so that pairs of 'off-the-shelf' lenses could be combined to produce a compound lens with any desired focal length. This allowed the f number of the system to be maximised by matching the size of the Fourier transform to the clear aperture of the BSO crystal. This optimisation of focal length was beneficial as it minimised the angular selectivity effects; and also allowed accurate matching of the scales of the Fourier transforms on the write and read-sides.

An important new effect, identified in this research, was that caused by the holographic grating vector becoming tilted with respect to the applied field. Experimental measurement, and a calculation of expected correlator fringe tilts, suggested that the magnitude of this effect on the write-side of the system could be comparable to that of the Bragg de-tuning effect on the read-side.

Taking all of the above factors into account, a design was produced that allowed the correlation of the specified images, at the desired rate. However, computer simulations predicted that the correlation plane intensities would only be between one and two orders of magnitude above the minimum sensitivity of a typical CCD camera. Thus, there was not much latitude for error in either predictions or practical realisation.

6.2 Extending the Design to Larger Images

It was discovered that the main equations governing the performance of the system could be expressed in terms of the coefficient N/r_{BSO} : where N was the number of pixels across the image; and r_{BSO} was the radius of the clear aperture of the BSO crystal. When this was done, some interesting predictions could be made regarding the extension of the system's performance. The main problem encountered was the angular selectivity of the hologram. This was exacerbated by a predicted increase in

'thick' behaviour as the holographic fringe spacing is decreased in order to deal with larger input images. This decrease is necessary in order to maintain the separation of the correlation plane and the undiffracted beam.

By examining the equations, it was found that three factors provided absolute limits on the value of N/r_{BSO} . These were: the need to keep the hologram's grating spacing greater than the critical value for the BSO; the need for the diffraction efficiency for light from the edge of the images to be non-zero; and the need to keep the longitudinal extent of the Fourier transform greater than the thickness of the BSO crystal.

For the correlation of two large images, the most severe limit on N/r_{BSO} was set by the angular selectivity of the hologram. It was shown that an upper limit would be the correlation of two 3000×3000 pixel images. However, the light utilisation of the system would be very poor with an effective diffraction efficiency of just 0.0002%. In addition to this, fringe tilt effects would reduce the response at the corners of the write image to 50% of the maximum. This would disrupt the form of the correlation peak; but there would be no simple way to alleviate the problem. It would therefore be difficult to realise such a system and a more practical upper limit might be the correlation of images of between 1000 and 2000 pixels across.

On the other hand, by utilising the relative lack of angular selectivity on the write-side, it would be theoretically feasible to correlate a 6000×6000 pixel image with a small reference template. Here, the limit on N/r_{BSO} was set by the depth of field in the Fourier plane. Importantly, the variations across the write-image, due to angular selectivity and fringe tilt, could be alleviated. To do this an intensity transmission filter would be placed in the correlation plane. The apparent diffraction efficiency would be reduced to the minimum value encountered for any point in the write image.

In either of the above systems, the optical path lengths would be measured in meters. Furthermore, large aperture, thin BSO crystals – $\sim 20 \times 20 \times 0.2\text{mm}$ – would be needed; as well as applied fields of around 10kVcm^{-1} .

6.3 Digital Simulation of SLM Performance

Software was written for the simulation of a variety of image distortions that can be produced in optically addressed SLMs. Both the phase and amplitude of the input images were considered. Because the input and output formats of the various programs were standardised, it was easy to apply a number of different distortions to the same image. This allowed a proper scientific investigation, into the relative effects of a variety of SLM faults, for any optical image processing task. The results produced, highlighted some shortcomings of published work in this field.

The 2^N factorial method was chosen as the experimental procedure around which the work was based. The advantage of this method was that it minimised the number of separate experiments that needed to be performed. At the same time, it allowed the direct effects of different distortions to be assessed when a number of them were affecting the SLM simultaneously. As well as their direct effects, the interactions of the various distortions can also be evaluated.

In an example experiment, a correlation task was studied. This involved the location of a small sub-feature within an image of the face of a generic, machined-metal component. It was shown how the experimental process could be fully automated; provided that the distortions were not so severe as to completely destroy the correlation peak. The experiment investigated three different distortions; using 256×256 pixel images. Eight combinations were possible, with each of the distortions either present or absent. For each combination, the correlation results were averaged over 12 images of the machined component at different positions in the input field. Thus, a total of 96 correlations were performed. These took about half an hour of computer run-time; using a Sun SPARC-station-2. If 512×512 images had been used, the experiment would have taken four times as long. If desired, six different factors could be investigated, using 512×512 images, in about 13 hours of run-time. This again assumes that, for each of the 64 possible combinations of the 6 factors, 12 correlations were performed. Thus it can be seen that, even without specialised hardware, this experimental technique is readily extendible.

6.4 Recommendations for Future Work

In the research reported here, the designed optical system could not be constructed because of an unexpected failure to acquire SLMs of the necessary quality. There is no doubt that a lack of guaranteed, high quality, commercially available SLMs is a limiting factor in the exploitation of optical technology. One of the reasons BSO was selected for this research was its high resolution as a holographic storage medium. If no SLMs can be found to allow high resolution images to be relayed to the optical system, then this advantage can not be exploited. Over recent years, there have been great advances in the processing speeds of digital Fourier transform based systems. A recently reported correlator [6.1] was claimed to run at 15 frames a second with 256×256 pixel images. To demonstrate a significant advantage over this and future digital processors, optical correlators would have to work with either very large images or at very high processing rates.

It has been demonstrated that BSO could, theoretically, be used to cross-correlate images of up to 3000×3000 pixels. The read-image could be updated at a rate limited solely by the data input and output to and from the system. However, the use of BSO for this task does not compare favourably with a new material; genetically engineered bacterio rhodopsin [6.2]. Unlike BSO, this material does not require an applied electric field. More importantly, useful diffraction efficiencies can be obtained with relatively thin films. These two facts remove all of the problems associated with using high f number lenses in BSO, as well as the constraints on input image position. Thus, it would allow the construction of much more compact architectures. This material would therefore be preferred to BSO for the correlation of large images.

For the cross-correlation of large images with small reference templates, an alternative to BSO is the use of a phase modulating SLM, as discussed in the introduction. In this case, the system would only require one laser. Since no hologram is formed, the requirement for interferometric stability would also be removed. The reduced complexity of these systems could make them more attractive than BSO for this purpose. As previously mentioned, the filters that could be placed in the Fourier

plane would not be limited to the transforms of real-valued inputs. This would give the system greater flexibility than the corresponding BSO-based correlator. The only drawback to this approach is the need to produce the transforms for the filters digitally. However, this would probably be done off-line, and a library of filters stored instead of a library of reference images. If binary phase only filters are used, the data storage requirements are reduced by a factor of 8 times; compared to standard grey-level images.

The most successful piece of work to come out of this research was that on the digital simulation methods for predicting SLM performance. There can be little doubt that BSO-based correlators are being overtaken by other, more attractive optical technologies. However, many future optical image processing systems will still depend on a spatial light modulator; either in the Fourier plane, in the image plane, or in both. The development of these devices can be both expensive and time consuming; as experiences during this project showed. It will therefore be of paramount importance that any devices produced, while being of sufficient quality for a particular image processing task, are not prohibitively expensive as a result of being over-engineered. At the same time, much research time can be saved by making sure that an SLM chosen to implement a particular system is in fact capable of reproducing images to the required standard. The digital simulation work reported here could be used to produce a specification/evaluation tool for various generic SLM types. This would be of use to both designers and users of SLMs and would ideally be employed at the initial stages of a project.

References: Chapter 6

- 6.1 Hose, M., "FFT Chips for Transform-Based Image Processing", Advanced Imaging, June 1992.
- 6.2 Hampp, N., Braüchle, C., Oesterhelt, D., "Bacteriorhodopsin and its Genetic Variants as Media for Dynamic Holographic Information Processing" Proc.IEE, Conf.Pub.342, Holographic Systems and Applications, p.35 (1991).

Appendix A

Operations Required to Perform Correlation

When computing the number of floating point operations it should be noted that on computers such as the Sun SPARC station it was found that the times taken to perform floating point addition and multiplication were identical.

In the first case, consider the image space, mask stepping approach. It should be noted that, using this technique, the images being correlated do not have to comprise the same number of pixels. The reference image will be assumed to be the smaller of the two and be $n_r \times n_r$ pixels. The edge enhancement will be done by stepping a mask across this image since it would be computationally inefficient to apply the edge filter to the larger image (an identical result is obtained either way). This edge enhancement mask will be $n_m \times n_m$ pixels. Finally, the object image is $n_o \times n_o$. First the edge enhancement mask is stepped over the reference image. For every point in the reference image $n_m \times n_m$ multiplications and $n_m \times n_m$ additions are performed. (If any part of the mask overlaps an edge of the image it is assumed that it wraps round to the opposite edge). Thus, a total of $n_r^2 n_m^2$ multiplications and additions are performed. After this the filtered reference image is stepped over the object image and multiplications and additions performed at each point. It should be noted that it is expected that both images and the edge enhancement filter would ideally be real valued so the additions and multiplications of the pixel values require only one floating point operation. The total number of multiplications and additions in the entire correlation process is then:

$$2n_r^2(n_m^2 + n_o^2) \approx 2n_r^2n_o^2 \quad (\text{A.1})$$

where, it has been noted that the number of pixels in a typical image space edge enhancement mask is usually much less than that in the object image.

For the Fourier transform approach, both the reference and object image are assumed to have the same number of pixels, $N \times N$. In order to compute the two dimensional Fourier transforms of the object and reference image, a series of one-dimensional Fast Fourier Transforms can be applied; first along the rows of the images and then down the columns [1.3]. Since the images are initially real valued, the N pixels along the row can be treated as an array of $N/2$ points [A.1]. However, the results of these transformations will be complex numbers. In order to perform a one dimensional FFT on an array of length N requires $(N \log_2(N))/2$ complex multiplications and $N \log_2(N)$ complex additions. So, transforming a real image requires N FFTs to be performed on arrays of apparent length $N/2$ and then N FFTs of arrays of length N . Thus, in order to transform the object and reference images the number of complex multiplications and additions are:

$$2 \times \left(\frac{N^2}{4} (\log_2(N) - 1) + \frac{N^2 \log_2(N)}{2} \right) = \frac{N^2}{2} (3 \log_2(N) - 1) \quad (\text{A.2})$$

and,

$$2 \times \left(\frac{N^2}{2} (\log_2(N) - 1) + N^2 \log_2(N) \right) = N^2 (3 \log_2(N) - 1) \quad (\text{A.3})$$

After taking these two transforms they must be multiplied together which requires N^2 complex multiplications and then must be multiplied by the Fourier transform of the edge enhancement filter which again is assumed to be real and thus requires N^2 real.complex multiplications. Finally the inverse transform of the result must be computed and this requires $2N((N \log_2(N))/2)$ complex multiplications and $2N(N \log_2(N))$ complex additions.

Finally, it should be remembered that a complex multiplication requires four floating point multiplications, a real.complex multiplication requires two, and a complex addition requires two floating point additions. The total number of floating point operations (multiplications and additions) is then found to be:

$$2N^2(3\log_2(N) - 1) + 2N^2(3\log_2(N) - 1) + 4N^2 + 2N^2 + 4N^2\log_2(N) + 4N^2\log_2(N) \quad (\text{A.4})$$

which is equal to,

$$N^2(20\log_2(N) + 2) \approx 20N^2\log_2(N) \quad (\text{A.5})$$

since for a 16×16 pixel image $20\log_2(N) = 80 \gg 2$.

Thus, to a first approximation if an $N \times N$ object image is assumed (n_o is set equal to N in the equation for the number of operations for correlation in image space) the size of reference image at which it becomes more economical to work in Fourier space rather than image space is given by:

$$20N^2\log_2(N) = 2N^2n_r^2 \Rightarrow n_r = \sqrt{10\log_2(N)} \quad (\text{A.6})$$

For example for a 512×512 image it is advantageous to use Fourier space for correlations with reference images of greater than 9×9 pixels.

